
**CELL MODELING**

THESIS

Matthew Lucas Campbell, Second Lieutenant, USAF

AFIT/GAM/ENC/02M-01

**DEPARTMENT OF THE AIR FORCE**
**AIR UNIVERSITY**

# AIR FORCE INSTITUTE OF TECHNOLOGY

**Wright-Patterson Air Force Base, Ohio**

AFIT/GAM/ENC/02M-01

# Cell Modeling

THESIS

Presented to the Faculty

Department of Mathematics and Statistics

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Applied Mathematical Analysis

Matthew Lucas Campbell, B.S.Mathematics

Second Lieutenant,USAF

March 2002

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

AFIT/GAM/ENC/02M-01

# Cell Modeling

Matthew Lucas Campbell, B.S.Mathematics

Second Lieutenant,USAF

Approved:

| | |
|---|---|
| | 14 MAR 02 |
| Dr. Dennis Quinn<br>Thesis Advisor | Date |
| | 14 Mar 02 |
| Dr. Charles Bleckmann<br>Committee Member | Date |
| | 14 Mar 02 |
| Dr. Glen Perram<br>Committee Member | Date |

*Acknowledgements*

I would like to thank my thesis advisor Dr. Dennis Quinn for his caring, understanding, and motivational guidance. Over the last several months he helped me to develop an understanding of how to conduct research that led to the completion of this thesis. His help allowed me to become a more mature student and researcher, capable of extending my skills into fields outside the realm of mathematics.

I would also like to thank Dr. Glen Perram and Dr. Charles Bleckmann for the effort they put forth in reviewing my thesis and providing valuable feedback.

I would like to thank Ever and Dr. Mark Oxley for their help with formatting issues, some of which were very time consuming.

I would like to thank my friends and roommates for putting up with me during the process of writing this thesis.

Most of all, I would like to thank my new wife for her support through all my efforts. Without her love, I couldn't have done it.

Matthew Lucas Campbell

*Table of Contents*

## *List of Figures*

## *Abstract*

The Air Force is currently developing new products that incorporate a variety of chemicals which may come in contact with product users. To define which chemicals are dangerous to the user, toxicity studies have been performed. However, analysis of toxicity ultimately requires models of the exposed cellular systems. This thesis provides an introduction of how to model and analyze small and large cellular systems. Understanding the underlying behavior of small models and their relation to large systems will lead to a better understanding of how the Air Force should construct intracellular models to assist in future toxicology studies.

Developing analysis techniques to include steady state analysis through linearization, and then considering small reaction systems, such as the Brusselator and Schnackenberg models, led to a basic understanding of model behavior. This knowledge was applied to create new models in an effort to begin a transition from previously created models to the construction of models unique to the Air Force.

Sensitivity analyses performed on existing systems furthered research efforts by developing knowledge of how systems behave under various initial conditions and perturbations of uncertain constant parameters. Analysis displayed great sensitivity within some models. This analysis was applied to a new model to look for interesting behavior such as oscillatory convergence. The new model was then incorporated into a larger model to determine how its behavior changed with respect to changes in the larger model.

This knowledge of how small systems behave in relation to larger systems should help the Air Force to develop and analyze intracellular toxicology models.

# Cell Modeling

## *I. Introduction*

### *1.1 Overview*

"The development of new products within the Air Force always involves the use of a variety of chemicals that may be needed in the production, use, or maintenance of the product. Often, these chemicals are either novel or are being used in ways unique to the Air Force. In order to facilitate the development of these products, while maintaining the safety of Air Force personnel, rapid and accurate methods for determining the toxicity of these chemicals are needed. Due to the complexity of human biology, standard toxicological assays are time consuming and, due to the lack of adequate understanding of toxicological mechanisms, may be inaccurate at predicting human risk. Because of these limitations, the Air Force has committed to the development of new toxicity assays that will generate large amounts of data on gene expression and protein concentration...The analysis of this data ultimately requires models of the entire set of cellular pathways on a quantitative basis which will ultimately reveal important feedback mechanisms that are undetectable without this comprehensive, quantitative approach." [25]

Cell models must be incorporated in understanding how the multiple metabolic interactions inside a cell lead to changes at the gene and protein level. Various research papers have been published to model the observed functions of cells. Over the last half of the twentieth century, mathematics has also played a large role in cellular modeling. "There has grown a sizable literature on the theoretical analysis of biological regulation, mostly by means of mathematical models." [15] As more information is gained about the internal processes of cells, mathematical models of

intracellular reactions have been incorporated to help keep track of these cellular processes.

Applications of mathematical models to cell processes are needed for many reasons. They are needed to identify the design principles of biochemical-based logic within a cell. This is key in understanding how normal and mutant cells behave in response to external and internal stimuli. More importantly, they can be incorporated into simulations that will give an understanding of how cells behave in a hypothesized environment. These simulations, combined with mathematical models, will predict effects of different nutrient concentrations or cell mutations on regulatory outcomes of the cell. This will lead to an improved understanding of cell toxicity by keeping track of each reaction and determining how the cell responds.

The cellular processes within human cells have great sophistication and are extremely complicated. For this reason, the modeling process of the human cell must be scaled down into simplified models that are more suitable for an initial investigation.

The cell is comprised of many complex systems that control cellular functions, such as, metabolic pathways, transcription, and translation. This thesis reviews work previously done on modeling transcription, translation, and metabolism as well as theoretical models already in existence such as the Brusselator and Schnackenberg's model. Mathematically modeling small systems and applying them to larger cellular processes will aid in understanding how cells are affected by toxic chemicals.

## *1.2 Problem*

In order for the Air Force to develop mathematical models to analyze cellular data, it is important to understand how small models of cellular activity were created, as well as, what behavior they displayed under various conditions. By constructing models of cellular activity, and evaluating their behavior in response to a sensitivity analysis, larger models may be better understood. Understanding the underlying

behavior of small models and their relation to large systems will lead to a better understanding of how the Air Force should construct intracellular models to assist in future toxicology studies.

### *1.3 Scope*

This research began by evaluating cellular models, such as the Hasty model [14], and hypothetical chemical models, such as the Brusselator [27] and Schnackenberg [30] models. The mathematical models presented in this thesis were derived deterministically from reaction equations that were based on mass action kinetics. Each reaction equation in the system represents complex functions of molecular interaction. Reactants within each equation were assumed to be either individual molecules or groups of molecules. The rate constants within the reaction equations reflect an average rate of reaction. The models are considered to be components of larger cellular systems.

### *1.4 Summary of Thesis*

This thesis is organized as follows:

Chapter 2 gives a brief overview of cellular components and functions. It then discusses previous and current work on the development of mathematical models of cellular activity. It introduces computer algorithms used to describe these and larger models. It also introduces hypothetical chemical oscillators to be reviewed as a part of this thesis in Chapters 3 and 4.

Chapter 3 presents techniques for describing the behavior of linear and nonlinear systems. It then derives and evaluates the behavior of current models. Finally, two new models are presented as examples of how current models can be extended.

Chapter 4 analyzes the behavioral changes of current and new models. It then addresses how some types of behavior may cause interesting outcomes within a larger cellular model. Finally, a small model is incorporated into a larger model

in an effort to produce a greater understanding of how large systems are affected by small component models.

Chapter 5 summarizes the work completed. It also presents conclusions reached, and gives recommendations for future work.

# *II. Background*

## *2.1 Overview*

Metabolic pathways can be described both qualitatively and quantitatively. Experimental observation can characterize metabolic pathways, but a mathematical approach can also uncover unknown facts about the process. For example, mathematical modeling may be used to evaluate concentration levels of reactants within a chemical or biological system where experiments become too expensive, dangerous, or time consuming to carry out. Together these two types of descriptions may uncover unaccounted-for or misunderstood processes leading to a better understanding of the cell's processes. Thus, learning about cellular processes from both experimental and mathematical modeling is a very beneficial way to gain understanding about those cellular processes.

Metabolic regulation is an important cellular process typically modeled with nonlinear ordinary differential equations (ODE's) based on chemical kinetics. The need for metabolic regulation arises as a result of different stimuli including stress, changes in the environment, and various nutrients ingested into the cell. Metabolic regulation can be referred to as signal transmission which occur through chemical interactions, enzymatic reactions, protein degradation, and production of intracellular messengers. It addresses questions such as 'How do pathways change to increase production of products?' and 'How does the cell respond to changes in nutrients?'[20] These pathways are linked through feedback loops, which use various regulators acting as messengers that are produced by one pathway and used as inputs for other pathways. Ideally, these pathways should link together to form large systems of ODE's that describe the cellular process as a whole. "The intricacy and variety of biological signaling networks often defy analysis based on intuition. Given these uncertainties, models such as these should not be considered as definitive descriptions of networks within the cell, but rather as one approach that allows us to understand

the capabilities of complex systems and devise experiments to test these capabilities [4]."

## *2.2 Biology*

Understanding the correct approach to modeling cellular behavior results from understanding the cellular mechanism being modeled. In order to model transcription and translation of a prokaryotic cell, a basic understanding of the biology background behind the cell must be outlined. This section contains quotes that address the basic definitions for eukaryotic and prokaryotic organisms, proteins and enzymes, DNA, RNA polymerase, mRNA, tRNA, rRNA, transcription, translation, operons, activators, repressors, and metabolic pathways along with some explanation of their functions. Also included are references where more detail can be found. Once these basic definitions are addressed, mathematical models of some of these biological systems will be considered.

*2.2.1 Bacteriophage λ.* Bacteriophage λ is a virus infecting *E. coli.* The bacteriophage λ life cycle is as follows. "The infection of the bacterial host *E. coli* begins when the virus specifically adsorbs to the cell and injects its DNA. The linear DNA then circularizes and commences directing the infection process. In the lysogenic mode, the phage DNA is stably integrated at a specific site in the host chromosome and so that it is passively replicated with the bacterial cell. Alternatively, the phage may take up the lytic mode in which the DNA directs its own replication, as well as the synthesis of viral proteins so as to result in the lysis of the host cell with the release of $\sim 100$ progeny phages. DNA damage, as is caused, for example, by UV radiation, induces the excision of the prophage DNA from the lysogenic bacterial chromosome and causes the phage to take up the lytic mode." [34:p 1090] More information about bacteriophage λ can be found in [34:p 1090] and [19:pp 83, 92, 335-336].

*2.2.2 Prokaryotic Organisms.* Prokaryotes are "organisms, usually bacteria, that have neither a membrane-bound nucleus enclosing their chromosomes nor functional organelles such as mitochondria and chloroplasts." [10:p 651] More information about prokaryotic organisms can be found in [6:p 502], [17:pp 25, 27, 52], [1:pp 20-21], [10:ch 1,2], and [34].

*2.2.3 Eukaryotic Organisms.* "Eukaryotes are defined by the division of each cell into a nucleus that contains the genetic material, surrounded by a cytoplasm, which in turn is bounded by the plasma membrane that marks the periphery of the cell. The cytoplasm contains other discrete compartments, also bounded by membranes." [17:p 27] More information about eukaryotic organisms can be found in [6:p 6], [1:p 237], [10:ch 1, 2], and [34].

*2.2.4 Amino Acids.* "Amino acids are the building blocks of proteins." [10:p 624] "In present-day living cells, large polypeptides-known as proteins-and polynucleotides-in the form of both ribonucleic acids (RNA) and deoxyribonucleic acids (DNA)-are commonly viewed as the most important constituents. A restricted set of 20 amino acids constitute the universal building blocks of the proteins, while RNA and DNA molecules are constructed from just four types of nucleotides each." [1:p 4-5] More information about amino acids can be found in [6:pp 68-70], [17:pp 8-11], [1:p 46], [10:ch 1, 2], and [34].

*2.2.5 Polypeptide.* A polypeptide is "a linear series of amino acids linked together with peptide bonds, it is also called a protein or protein chain." [10:p 648] More information about polypeptides can be found in [6:pp 68-76], [17:pp 3-25], [1:pp 107-108, 111-135], [10:ch 1,2], and [34].

*2.2.6 Metabolic Pathways.* "As a whole, metabolism is concerned with managing the material and energy resources of the cell. Some metabolic pathways release energy by breaking down complex molecules to simpler compounds. These

degenerative processes are called catabolic pathways. A major thoroughfare of catabolism is cellular respiration, in which the sugar glucose and other organic fuels are broken down to carbon dioxide and water. Energy that was stored in the organic molecules becomes available to do the work of the cell. Anabolic pathways, in contrast, consume energy to build complicated molecules from simpler ones. An example of anabolism is the synthesis of a protein from amino acids. Catabolic and anabolic pathways are the downhill and uphill avenues of the metabolic map." [6:pp 83-84] More information about metabolic pathways can be found in [1:p 87], [10:ch 1, 2], and [34].

*2.2.7 DNA.* "Biological instructions are encoded in the molecules known as DNA(deoxyribonucleic acid). DNA is the substance of genes, the units of inheritance that transmit information from parents to offspring." [6:pp 6, 77] More information about DNA can be found in [6:p 77], [17:p 81], [1:pp 4-5], [10:ch 1, 2], and [34].

*2.2.8 RNA Polymerase.* "RNA polymerase, the enzyme responsible for the DNA-directed synthesis of RNA, was discovered independently in 1960 by Samual Weiss and Jerard Hurwitz...All cells contain RNA polymerase. In bacteria, one species of this enzyme synthesizes all of the cell's RNA except the short RNA primers employed in DNA replication....RNA polymerase, which has a characteristic large size, binds to DNA as a protomer. This large size is presumably a consequence of the haloenzyme's several complex functions including template binding, RNA chain initiation, chain elongation, and chain termination...Once an RNA polymerase molecule has initiated transcription and moved away from the promoter, another RNA polymerase can follow suit. The synthesis of RNAs that are needed in large quantities, ribosomal RNAs, for example, is initiated as often as is sterically possible, about once per second." [34:pp 919-920, 925] More information about RNA polymerase can be found in [6:p 300], [17:pp 811, 824], [1:p 252], [10:ch 1, 2], and [34].

*2.2.9 Transcription.* "Transcription generates a single-stranded RNA identical in sequence with one of the strands of the duplex DNA. Several different types of RNA are generated by transcription; the three principal classes involved in the synthesis of proteins are: messenger RNA (mRNA), transfer RNA (tRNA), and ribosomal RNA (rRNA)." [17:p 154] "In a metabolically active cell, about 3-5% of the cellular RNA is mRNA, about 90% is rRNA, and about 4% is tRNA." [10:p 30] Furthermore, "the in vivo rate of transcription is 20 to 50 nucleotides per second at 37° C as indicated by the rate at *E. coli* incorporate H-labeled nucleosides into RNA." [34:p 925] More information about transcription can be found in [6:pp 296-297, 300-302, 315], [1:p 366], [10:ch 1, 2], and [34].

*2.2.10 Operator.* The operator is "the region of DNA that is upstream from a prokaryotic gene and to which a repressor or activator binds." [10:p 647] More information about operators can be found in [6:p 338], [17:p 166], [1:p 417], [10:ch 1, 2], and [34].

*2.2.11 Effector Molecule.* An effector molecule is "a low-molecular-weight compound that modifies the function of a regulatory protein." [10:p 634] More information about effector molecules can be found in [10:ch 1, 2].

*2.2.12 Activator.* An activator is: "(1) A substance or physical agent that stimulates transcription of a specific gene or operon. (2) A protein that binds to an operator and enhances the rate of transcription. Also called activator protein." [10:p 623] More information about activators can be found in [6:pp 354-355], [17:p 359], [10:ch 1, 2], and [34].

*2.2.13 Activator site.* An activator site is "a DNA sequence to which an activator protein binds. Also called activating site." [10:p 623] More information about activator sites can be found in [10:ch 1, 2].

*2.2.14 Repressor.* A repressor is "a protein that binds to the operator or promoter region of a gene and prevents transcription by blocking the binding of RNA polymerase." [10:p 653] More information about repressors can be found in [6:p 338], [17:p 339], [1:p 249], [10:ch 1, 2], and [34].

*2.2.15 Operons.* "All together, the operator, the promoter, and the genes they control-the entire stretch of DNA required for enzyme production for the tryptophan pathway-is called an operon." [6:p 338] More information about operons can be found in [6:p 338], [17:p 166], [1:p 417], [10:ch 1, 2], and [34].

*2.2.16 Messenger RNA.* "Each gene along the length of the DNA molecule directs the synthesis of a type of RNA called messenger RNA (mRNA). The mRNA molecule then interacts with the cell's protein-synthesizing machinery to direct the production of a polypeptide. We can summarize the flow of genetic information as DNA→RNA→protein." [6:p 77] More information about messenger RNA can be found in [6:p 296], [17:p 717], [1:p 223], [10:ch 1, 2], and [34].

*2.2.17 Ribosomal RNA.* "There are two major types of rRNA. The larger of these rRNA's combines with a set of proteins to form a ribonucleoprotein complex called the large ribosomal subunit. The smaller rRNA combines with another set of proteins to form a smaller ribosomal subunit. During protein synthesis, one large ribosomal subunit and one small ribosomal subunit combine to form a ribosome." [10:p 30] Furthermore, "even *E. coli* needs seven copies of its rRNA genes to keep up with the cell's need for ribosomes." [1:p 378] More information about ribosomal RNA can be found in [6:p 306], [17:p 179], [10:ch 1, 2], and [34].

*2.2.18 Ribosome.* "The actual sites of protein synthesis are cellular structures called ribosomes." [6:p 77] More information about Ribosomes can be found in [6:pp 111, 306, 506, 523], [17:pp 6, 33, 159-162], [1:p 223], [10:ch 2], and [34].

*2.2.19 Transfer RNA.* "In the process of translation a cell interprets a genetic message and builds a protein accordingly. The message is a series of codons along an mRNA molecule, and the interpreter is transfer RNA (tRNA). The function of tRNA is to transfer amino acids from the cytoplasm's amino acid pool to a ribosome." [6:p 304] More information about transfer RNA can be found in [17:pp 153, 213], [1:pp 227-230], [10:ch 1, 2], and [34].

*2.2.20 Translation.* "Translation converts the nucleotide sequence of RNA into the sequence of amino acids comprising a protein. An mRNA is translated into a protein sequence: tRNA and rRNA provide other components of the apparatus for protein synthesis. The entire length of an mRNA is not translated, but each mRNA contains at least one coding region that is related to a protein sequence by the genetic code: each nucleotide triplet (codon) of the coding region represents one amino acid." [17:p 154] More information about translation can be found in [6:pp 296-297, 304], [1:p 25], [10:ch 1, 2], and [34].

## *2.3 E-Cell*

The E-Cell project began in 1996 at Keio University [33]. It's goal is to model various biochemical processes with the intended goal of eventually modeling the entire cell. E-Cell's main attraction is its ability to integrate metabolic pathways with high-order cellular processes. These processes include protein synthesis through transcription and translation and membrane transport.

E-Cell is run under a UNIX, Linux, MSDOS, or MS Windows operating system. It requires the user to define the cell's molecules, locations, and concentrations. It then requires the user to input the set of reaction rules that govern all these processes. E-Cell then computes the concentration and location of each molecule at each time increment. Furthermore, in UNIX or Linux, a graphical interface allows the user to monitor the cell as each reaction occurs.

Using E-Cell, Tomita[33] et al. were able to construct an electronic cell with 127 genes that was capable of "self-support." The cell model included "genes for transcription, translation, glycolysis pathways for energy production, membrane transport, and the phospholipid biosynthesis pathway for membrane structure." [33]

### *2.4 McAdams, Arkin, Shapiro, and Bhalla*

Biochemical and genetic approaches have identified a number of the molecular mechanisms of bacteria. A problem attracting considerable research effort is the detailed understanding of the dynamic behavior of large systems of genes and related proteins and how they interact to control cellular functions over the cell's life cycle. Harley McAdams and Lucy Shapiro[22] investigated this problem through the use of genetic networks. These genetic networks attempt to incorporate some of the techniques used in electrical circuits to gain a better understanding of intracellular mechanisms. Genetic networks consist of hundreds of genes and consequently are complex. "Extension of metabolic modeling methods to include more realistic genetic regulatory mechanisms is a current challenge to the field." [20] For instance, the coupled reactions controlling cell division in prokaryotic cells have not been identified. Due to the commonalities in the function of these biochemically based genetic circuits and electrical circuits [20], a new modeling technique has emerged as a way of integrating conventional biochemical kinetic modeling within the framework of a circuit simulation.

In genetic networks, the protein production encoded by one gene often regulates expression of other genes. Activation depends on the level of expression of one gene required to control the expression of other genes. The concept of genetic circuits was motivated by electrical circuits, but they have important differences. McAdams and Shapiro state that, "Common transistor circuits can operate at more than $10^8$ cycles per second. In contrast, the protein signal-controlled switching rate in genetic circuits is around $10^{-2}$ per second." [22] Due to this, McAdams and

Shapiro incorporated time delays into their system. They preformed simulations of gene expression which resulted in the proteins in the simulation being produced in short bursts and in variable numbers occurring at random times. As a result, they claim there can be large gaps in time between successive events in regulatory changes across a cell. Additionally, randomized patterns of expression of competitive effector molecules can produce outcomes in switching mechanisms that select between alternative regulatory paths. As a result of following different paths, different types of cells may occur. McAdams and Shapiro observed this in their model of the Bacteriophage $\lambda$ lysis-lysogeny decision circuit.

Molecules that control gene expression, such as activators and repressors, may act at extremely low intracellular concentrations. For this reason, large fluctuations in reaction rates may occur, causing variations of concentrations of each molecular species. In spite of this, many regulatory pathways in cells have highly predictable outcomes. To achieve regulatory reliability, cells use redundant genes and networks, as well as, feedback loops to buffer pathways against mutational or environmental perturbations.[21]

Bhalla and Iyengar[4] have also studied metabolic pathways using circuit components. They have shown that many distinct signaling pathways allow the cell to receive information, process it, and respond. Their results include networks that exhibit properties such as integration of signals across multiple time scales, generation of distinct outputs depending on input strength and duration, and self-sustaining feedback loops. They state that, "feedback loops can result in bistable behavior with discrete steady-state activities, well-defined input thresholds for transition between states and prolonged signal output, and signal modulation in response to transient stimuli." [4]

## *2.5 Smolen et al. and Hasty et al.*

*2.5.1 Smolen et al.* Smolen et al. [32] investigate the possibility that multi-stability and oscillatory behavior occur in genetic regulatory systems. They created kinetic models that incorporate known features of genetic regulatory systems. "These include autoregulation and stimulus-dependent phosphorylation of transcription factors (TFs), dimerization of TFs, crosstalk, and feedback. The simplest model manifested multiple stable steady states, and brief perturbations could switch the model between these states. ... In slightly more complex models, oscillatory regimes were identified." [32]

Smolen et al. looked at different complexities of models to determine where strange behavior in genetic systems may occur. They began by considering "a relatively simple model of transcription factors (TFs) subject to positive and negative autoregulation of their own transcription." [32] First, they considered "signal-transduction pathways in which stimuli lead to second messenger generation and phosphorylation of TF's, which in turn bind to DNA sequences known as responsive elements and thereby regulate the transcription of specific genes. The regulatory activity of TFs is often modulated by phosphorylation and by intermolecular interactions. For example, TFs often bind to DNA as homodimers or as hetrodimers of different TF family members...As a result, some TFs, such as Jun, autoregulate their own transcription [24]...Thus responsive elements affecting TF gene transcription can provide crosstalk and positive feedback." [32] Smolen et al. made some simplifications to the model so it could be appreciated intuitively and described mathematically. "A single transcriptional activator, which we term TF-A, is considered as part of a pathway mediating a cellular response to a stimulus. The TF forms a homodimer that can bind to responsive elements (TF-REs). The tf-a gene incorporate one of these responsive elements, and when homodimers bind to this element TF-A transcription is increased. Binding to the TF-REs is independent of dimer phosphorylation. Only phosphorylated dimers, however, can activate tran-

scription. The fraction of dimers phosphorylated is dependent on the activity of kinases and phosphatases whose activity can be regulated by external signals. Thus this model incorporate both signal-activated transcription and positive feedback on the rate of TF synthesis." [32]

Smolen et al. expanded their model by incorporating both positive and negative feedback. "Responsive elements have also been found that regulate genes for potent transcriptional inhibitors...Such responsive elements provide negative feedback loops. Given the above interactions, there is the possibility for rich dynamic activity...However, models similar to the one above have only one type of feedback: either positive or negative. Without both types of feedback, such a model cannot be expected to support oscillations. For example, positive feedback can act to drive [TF-A] to high levels, but then there is no process to bring [TF-A] back down. To investigate the effects of negative feedback, we introduced a protein, TF-R, that represses transcription by binding to the TF-REs. Its rate of synthesis is increased by binding of the TF-A dimer to a TF-RE. TF-R competitively inhibits binding of TF-A dimers to TF-REs; thus it inhibits the transcription of the genes tf-a and tf-r." [32]

Using a bifurcation graph that plotted steady state concentrations of the transcription factor as a function of a parameter, Smolen et al. determined that depending on different parameters, the steady state concentrations of this model may exhibit strange behavior in the form of two stable steady states and one unstable steady state. This behavior is known as bistability. Smolen et al. stated that bistable "transitions could correspond physiologically to brief stimuli, such as exposure to hormone, leading to long-lasting increases or decreases in the levels of particular proteins." [32]

*2.5.2 Hasty et al.* Hasty et al.[14] based their work, in part, on that of Smolen et al. [32] In their analysis, the original model created by Smolen et al.

that included only the activator was incorporated into the following system which resembles the stoichiometric equations of chemistry.

$$2X \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} X_2 \tag{2.1}$$

$$D + X_2 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} DX_2 \tag{2.2}$$

$$DX_2 + P \xrightarrow{k_3} DX_2 + nX \tag{2.3}$$

$$X \xrightarrow{k_4} A \tag{2.4}$$

In the presentation by Hasty et al., they assume the following. "Let X denote the protein and D a specific binding site on a DNA promoter. The production of X takes place when the dimer $X_2$ forms a complex with the DNA at site D...The first reaction represents the dimerization of X, and the second the formation of a DNA-promoter complex. In the presence of RNA polymerase P and the reactants required for transcription, this complex leads to the effective transcription of the gene coding for X. The produced mRNA is then translated into protein X, where n is the number of proteins per transcript. Both transcription and translation are represented by the third equation. Lastly X degrades via reaction four...Further, the first two reactions are orders of magnitude faster than transcription and degradation." [14]

Analysis led Hasty et al. to conclude that "bistability arises as a consequence of both the dimerization of the protein, and the competition between the production via transcription and the degradation in the last reaction." [14]

In the second paper by Hasty et al.[13] they model the effects the bacteriophage $\lambda$ virus by analyzing the evolution of the $\lambda$ repressor protein. "Bacteriophage $\lambda$ has two alternate life cycles, lytic and lysogenic growth." [19:83, 92, 335-336] "In the context of the lysis-lysogeny pathway in the $\lambda$ virus, the autoregulation of $\lambda$ repressor

expression is well characterized [16], [28]... We present two models describing the regulation of such a network." [13]

Hasty et al. constructed the model to represent the different effects bistability has on mutant and normal operator regions of the bacteriophage $\lambda$ virus. "Although the full promoter region in $\lambda$ phage contains the three operator sites known as OR1, OR2, and OR3, we first consider a mutant system whereby the operator site OR1 is absent from the region. The basic dynamical properties of this network, along with a categorization of the biochemical reactions, are as follows [16], [28]. The gene cI expresses repressor (CI), which in turn dimerizes and binds to the DNA as a transcription factor. In the mutant system, this binding can take place at one of the two binding sites, OR2 or OR3. (Here we ignore nonspecific binding.) Binding at OR2 enhances transcription, which takes place downstream of OR3, whereas binding at OR3 represses transcription, effectively turning off production." [13]

The reactions describing the mutant operator region model are below.

$$2X \overset{k1}{\rightleftharpoons} X_2 \tag{2.5}$$

$$D + X_2 \overset{k2}{\rightleftharpoons} DX_2 \tag{2.6}$$

$$D + X_2 \overset{k3}{\rightleftharpoons} DX_2^* \tag{2.7}$$

$$DX_2 + X_2 \overset{k4}{\rightleftharpoons} DX_2X_2 \tag{2.8}$$

$$DX_2 + P \overset{kf}{\rightarrow} DX_2 + P + nX \tag{2.9}$$

$$X \overset{kd}{\rightarrow} A \tag{2.10}$$

Equation (2.5) of the new model involves the dimerization of the $\lambda$ repressor protein ($X$) to create a dimer protein. The dimer protein ($X_2$) can bind to the DNA ($D$) at one of three operator sites. The second operator site, labeled as ($DX_2$) in the model above, enhances transcription, which is consistent with lysogenic growth

of Bacteriophage $\lambda$. The third operator site, labeled as $(DX_2^*)$ in the model above, which exists down stream from the second operator site represses transcription, which is consistent with lytic growth of Bacteriophage $\lambda$. Also, a dimer protein can bond to the third site, and then the second $(DX_2X_2)$. In this case, the second site acts as a repressor and blocks the activator site from moving on. Finally, RNA polymerase joins the activated site to complete transcription, after which, translation is completed, and as a last process, the repressor protein degrades. These are much slower reactions that the first four and are also irreversible.

As in the previous model, Hasty et al. states that the bistability in this system arises from the competition between the production of the repressor protein seen in equation (2.9), its dimerization in equation (2.5), and its degradation from equation (2.10).

### *2.6 Chen et al.*

Chen, He, and Church[7] formulated a different approach to studying large systems of genes. They explain how to construct their model from initial data given by samples of mRNAs and proteins. Their results suggest that given accurate data, their models can determine most gene regulation accurately at the genome level.

They begin by considering the simplifications that must be made about the actual processes of transcription and translation in order to fit them into the model. They ignore direct feedback from mRNAs to genes. Instead they assumed all feedback to the gene is a direct result of proteins. Furthermore, they assumed "the translation mechanism is relatively stable (at least for a short time), so the feedback from proteins to mRNAs has no effect. Each mRNA and protein molecule degrades randomly, and its components are recycled in the cell." [7] Finally, they assume there is no feedback from the metabolic pathways to transcription. They point out that feedback from the metabolic pathway to transcription may actually have some

effect on the system. However, it is a simplification they have chosen to more easily solve their model.

After making their assumptions, they were able to mathematically describe their model through the nonlinear dynamic system below.

$$\begin{aligned} \frac{d\mathbf{r}}{dt} &= f(\mathbf{p}) - V\mathbf{r} \\ \frac{d\mathbf{p}}{dt} &= L\mathbf{r} - U\mathbf{p} \end{aligned} \tag{2.11}$$

The variables $\mathbf{r}$ and $\mathbf{p}$ are functions of time $t$. The quantity $n$ is the number of genes being included in (2.11). $\mathbf{r}$ is the $n$-dimensional vector valued function for mRNA concentrations. $\mathbf{p}$ is the $n$-dimensional vector valued function for protein concentrations. $f$ is the set of functions of proteins $\mathbf{p}$ that describe transcription. $L$ is the $n \times n$ dimensional matrix describing translation. $V$ is the $n \times n$ dimensional matrix describing degradation of mRNAs. Finally, $U$ is the $n \times n$ dimensional matrix describing degradation of proteins.

For equations (2.11), the following observations can be made. mRNA concentrations $\{\mathbf{r}\}$ increase as transcription $\{f(\mathbf{p})\}$ increases, and decreases as the degradation of mRNA $\{V\mathbf{r}\}$ increases. Protein concentrations $\{\mathbf{p}\}$ increase as translation of mRNA $\{L\mathbf{r}\}$ increases and decreases as degradation of proteins $\{U\mathbf{p}\}$ increases. Translation rates and degradation rates are assumed constant for both mRNA and proteins so $L$, $V$, and $U$ are diagonal matrices with no zeros on the diagonal. Finally, they assume that for this model, there is no time delay. They leave time delay for another model.

They solve their problem by linearizing the transcription functions, $f(\mathbf{p})$. The resulting linear system has a unique solution which can be written in terms of the eigenvalues and eigenvectors of the system.

Additional insight concerning this problem may be found by considering gene regulatory networks. Although genes act as regulators for other genes, this regulation often involves only a few number of regulators. Hence, all genes do not act as regulators for each other. Chen claims "that the number of regulators for a gene is small, mostly less than 10 [7]." Savageau adds validity to this claim by stating "molecular analysis of gene regulation in bacteria has shown that most gene circuits are governed by a small number of regulators usually one to three. In eukaryotes the numbers are larger in some cases, but seldom more than a dozen regulators influence a given gene circuit....The most biologically-suggestive behaviors were found when each circuit was subject to two or three regulatory interactions [29]." This fact allows for the production of a sparse matrix.

In discussing the limitations of their model (2.11), Chen et al. make note of specific potential flaws in the model. They point out that although the model does not include time delay, it does capture many features of gene expression. Furthermore, the main design flaw comes from the ignorance of other regulators such as those constructed via the metabolic pathway. Other potential hazards of this model include the assumption of periodic cell cycles and small numbers of regulators. Failure to meet both assumptions prevents the model from being reconstructed via the data.

### *2.7 Heinrich et al.*

Heinrich et al.[15] described metabolic systems through mathematics and chemical kinetics. They concluded that the number of hypothetical biological concepts required to explain data results may be scaled down by the use of models. They explained how to set up and simplify a model by only including key elements. Then,

how to construct the model, using equations based on thermodynamics and kinetic equations. They analyzed the dynamics behind these models, and identified principles of stability. Their rational for this analysis was to show how biological dynamics can be better understood when incorporated with mathematical models.

Heinrich et al. began by describing the basic need for mathematical modeling in the metabolic pathway. Modeling can be used to discover the fate of some of the intermediary proteins and enzymes in the metabolic pathway. These proteins and enzymes may behave differently under different concentrations. To understand these differences, models are needed to analyze the basic dynamics and stability regions of these intermediary protein and enzyme concentrations.

To deal with the complexity of biological entities, Heinrich et al. describe how to simplify models based on the models' processes. The rate at which regulatory properties of metabolism are expressed depend, in part, on the components within metabolism, concentration of the different elements, and the kinetics governing the reactions. It is important to understand which components are necessary to include in the model, and how they are connected. Heinrich et al. explained that "only the essential reacting variables have to be considered since faster variables are often in a quasi-steady state and slower ones are constant. ... Whole pathways may be substituted by single reactions in this way. For instance, Selkov [31] considered both hexokinase and phosphofructokinase as a single phosphorylating reaction and the phosphoglycerate kinase and pyruvate kinase reaction as one reaction regenerating ATP." [15] Many reactions are quite complicated, especially when enzymes are involved. Reactions may fluctuate either quickly or slowly depending on the concentration of the reactants. For this reason, simplified rate functions are often needed to describe the effects of the different concentrations of reactants. They are incorporated into systems of first order differential equations that represent the flux in concentrations of all the reactants being modeled. This flux is determined by the equations below.

$$\frac{dS_i}{dt} = \sum_{j=1}^{r} c_{ij} v_j \quad \text{where } i = (1, ..., n) \tag{2.12}$$

The variables in equation (2.12) are describes as follows: $S_i$ is the concentration of the reactant; $n$ is the number of reactants; $r$ is the number of reactions in the model; $v_j$ is the combination of different reactants associated with $S_i$; $c_{ij}$ is the number of molecules of reactant $S_i$ participating in reaction $j$; and $t$ is the time. Now, $c_{ij}$ is positive if reactant $S_i$ is a substrate of reaction $j$, where as $c_{ij}$ is negative if reactant $S_i$ is a product of reaction $j$. Furthermore, $v_j$ can be determined via the following equation.

$$v_j = v_j^- - v_j^+ = (k_j \underbrace{\prod_i S_i^{c_{ij}}}_{\text{products}} - k_{-j} \underbrace{\prod_i S_i^{c_{ij}}}_{\text{substrates}}) R_j(S_i, p_k) \tag{2.13}$$

Equation (2.13), $v_j^+$ is the forward reaction rate and $v_j^-$ is the reverse reaction rate. Also, $k_j$ is the forward reaction rate coefficient and $k_{-j}$ is the reverse reaction rate coefficient. $R_j$ is the function containing all the information related to special catalytic rate functions such as the Michaelis-Menton equation. Lastly, $p_k$ is the kinetic parameters associated with special catalytic rate functions such as $K_m$-values.

Equation (2.12) can be represented in the following abbreviated manner,

$$\frac{dS_i}{dt} = f_i(S_1, ..., S_n) \quad \text{where } (i = 1, ..., n) \tag{2.14}$$

where $f_i$ is the flux of reactant $S_i$ for $(i = 1, ..., n)$.

An analysis of this mathematical model can be performed to determine the dynamics behind the model, and identify principles of stability. The steady state of system (2.14) is achieved when $\frac{dS_i}{dt} = 0 = f_i(S_1^0, ..., S_n^0)$, where $S_i^0$defines the steady

state concentrations of the individual element $S_i$. A discussion of stability will be presented in Chapter 3 when specific applications are considered.

The domain surrounding a steady state point may be stable, or unstable. If a concentration is pulled away from a stable steady state, it will eventually move back to that steady state. If the concentration is pulled away from an unstable state, it continues to move away. Although biological organisms generally involve stable domains, in nonlinear systems, locally stable domains may be very small. Thus, a small perturbation from the stable state may allow it to return, but a larger perturbation may move it out of the local domain and in the direction of another steady state, a limit cycle, or cause it to diverge.

Small systems that can be evaluated have the capability to be analyzed via bifurcation diagrams in the form of plots that represent steady state concentration as a function of a parameter. If the parameter is changed, then the steady state concentrations may increase or decrease. Given enough change in the parameter, the steady state concentration may fork into multiple steady states such as a model by Baras et al. [3] shows, or become oscillatory such as Gillespie [9] point out though use of the Brusselator equation. This diagram is critical in analyzing dynamic systems because it can display changes of the stability of the system as parameters change.

To prove their case, Heinrich et al. [15] evaluated a small system without linearizing it. They obtained numerical solutions of the system. The system exhibited interesting behavior such as limit cycles and multiple stable states. This system was an approximate representation of a form of metabolic control called feedback loops. One variation of a feedback loop is where a product of a pathway is used as a rate controlling mechanism to control its own pathway. This type of feedback can cause the pathway to have multiple steady states. Thus, changes in concentrations may lead to different steady states and eventually to different biological behavior.

*2.8 The Brusselator and Schnakenberg equations*

The Brusselator model was created as "a scheme which may lead to a chemical symmetry-breaking instability." [27] It involves four reaction equations with two intermediate reacting components. One of the components acts as an autocatalyst. That is, it catalyzes its own formation. In doing so, it involves a tri-molecular reaction step. Initially introduced as a chemical oscillator, Prigogine and Lefever considered it "physically unrealistic because of the tri-molecular step." [27] However, it provides a basic understanding of how oscillation occurs, and may lead to a greater understanding of oscillations in biological systems when incorporated into more general systems of reaction equations.

Schnakenberg [30] considered an extremely simple mathematical model displaying oscillatory limit cycle behavior. He began his analysis by considering a model of general chemical reactions. Systems of chemical reactions consist of stoichiometric equations. Stoichiometric systems of equations take the form

$$A_j + c_{1j}S_1 + c_{2j}S_2 + c_{3j}S_3 + ... \rightleftharpoons B_j + c'_{1j}S_1 + c'_{2j}S_2 + c'_{3j}S_3 + ... \tag{2.15}$$

where $A$ and $B$ are fixed concentrations of external reactants, and all other variables are described in equation (2.12). To obtain the desired mix of simplicity in the model and interesting behavior in the solution, Schnackenberg chose to use two variables of reactants to create his model which is of the form of equation (2.14) with $n = 2$. Schnackenberg linearized his nonlinear system and solved to find the eigenvalues of the linearized system. This analysis will be presented in Chapter 3 when several examples are considered.

With the understanding that at least one unstable steady state point is needed to get oscillatory behavior, Schnackenberg continued the process of creating a simple oscillator. He did this by first limiting the number of equations he uses. Systems of four reactions, such as the Brussilator, have already been modeled, so to model

a system with more reactions than three wouldn't lead to a simplest oscillatory equation. Systems of only a single reaction can be ruled out because their steady state is the thermodynamic equilibrium. For this reason, oscillation cannot occur. Systems with two equations and two variables can also be ruled out because they are purely stable or unstable. Unstable steady states in this system will only create harmonic instability that degenerates out to infinity. Thus, a three reaction system is considered.

All three equations can be combined to produce the required conditions of a unstable steady state. Thus, completing conditions to produce a limit cycle with only three equations and two variables. In producing a limit cycle, it is important to be aware that the conditions applied to create a limit cycle don't rule out the possibility that both an unstable and a stable steady state lie within or outside the limit cycle. In some cases, all the conditions for limit cycles may exist, but if there is a stable node inside this region as well, the concentrations may approach that stable node instead of the limit cycle.

Limit cycles are of great importance in biology and particularly in metabolic pathways. They were discussed back in 1968 by E. Selkov [31] as models for self-oscillations in glycolysis. It is not unreasonable to think that such models may be extended to other systems in biology to include transcription and translation.

### *2.9 Gepasi 2*

Gepasi 2 is a free software package developed by Pedro Mendez[23] for the intended use of simulating models of biological and chemical reactions. Its development was based on biochemical systems theory created by Savageau[29], and metabolic control analysis explained by Heinrich et al. [15], as well as various pre-existing simulators. It has an easy to use interface and will run on MS-DOS and MS-Windows. It has no plotting capabilities, but has the capability to incorporate a graphics simulator. It's purpose is to mathematically simulate chemical systems

with both predefined and user defined kinetic equations. It comes with a short user manual as well as examples of simulations.

Gepasi 2 first requires the input of stoichiometric equations representing chemical reactions. It has the capability of incorporating up to 45 metabolites and 45 reactions. It then transforms these stoichiometric equations into ordinary differential equations (ODE) that can be solved using a ODE solver. Gepasi 2 uses the free solver LSODA developed by Petzold [26]. LSODA is a ordinary differential equations solver that can test whether or not a system is stiff by determining the ratio of the largest and smallest eigenvalues of the system. If the ratio is large in magnitude, then the system is stiff and solved using a stiff solver. If the ratio is small in magnitude, the system is non-stiff, a non-stiff solver can be used. LSODA solves stiff systems using the backwards differentiating formula. This is a implicit method whose region of absolute stability contains the entire negative real axis and a large portion of the complex plane so as to retain a stable solution. LSODA solves non-stiff systems using the Adams method. The Adams method is an explicit method with a smaller region of stability that involves time steps. It requires input from the user to define its time step size by asking for an end time and a number of intervals. [2:pp 407-414]

Other functions of Gepasi 2 include an ability to solve for steady states and to reduce the stoichiometric matrix defined by the user. It uses a damped Newton-Raphson method to solve for steady states of the system. As a precaution, this method ends after $10^{10}$ time units (in seconds, this would correspond to about 300 years). It is assumed that if the system doesn't converge at that point, a meaningful steady state cannot be reached. It reduces the stoichiometry matrix using the Gauss reduction method.

Gepasi 2 has 35 of the most commonly used kinetic equations predefined in the simulator so as to be easily integrated into the users system. These kinetic types include the Henri-Michaelis-Menton equation [34:ch 13] and the Hill equation

[34:pp 217-219] for both forward and reverse reactions. It also has the capability of implementing kinetic equations designed to act as feedback loops. These loops can represent either activator or inhibitors depending on the kinetic equation used. User defined kinetic types can also be implemented. Mendes[23] says that in cases where partial derivatives are input, they can be solved using finite differences.

## *2.10 Other Work*

*2.10.1 V-Cell.* V-Cell was developed by Leslie Loew and James Schaff of the University of Connecticut Health Center. Although not used in this thesis, it is important to address V-Cell due to its potential as a simulator. V-Cell is a simulator that has the capability to model the cell's shape, volume, and other physical features. It also models how molecules diffuse through the cell by using a system developed through the understanding of actual tests of diffusive activity in cells. Although V-Cell is a smaller simulator than E-Cell, it has the potential for creating greater realism within a cell.

*2.10.2 Global Identifiably Algorithms.* Once cellular functions are modeled and formulated as first order ordinary differential equations, they need to be solved. This would be an easily solved problem if the rate coefficients were known. Unfortunately, the rate coefficients are very hard to measure. Furthermore, it is often the case that small errors in the measurements lead to large errors in the solution, so the measurements must be extremely accurate.

Cobelli et al.[8] considered algorithms such as the Bunchberger algorithm to estimate the rate constants of nonlinear models. They incorporate the Bunchberger algorithm into a larger algorithm to check for global identifiably, a property that, under the assumptions of noise-free observations and error-free model structure, guarantees uniqueness of the solution.

## *2.11 Summary*

This Chapter described several examples of modeling transcription and translation and several examples of modeling metabolic pathways. Chapter 3 presents a detailed analysis of these models as well as an approach for analyzing large systems which include transcription, translation, and metabolism combined.

# *III. Analysis of Models of Transcription, Translation, and Metabolism*

## *3.1 Overview*

Models describing cellular activity have been created to predict cellular behavior for different concentrations of component molecules, and different rate constants. Many of these models are described using systems of reaction equations that can be reformulated into systems of differential equations that track continuous movement of each component though time. These systems of differential equations can provide detailed analysis about the state of the system. This chapter will begin by indicating how to linearize about a steady state and then determining local behavior from the eigenvalues of the linearized system. It will then describe models of equations that exhibit oscillatory behavior and multiple steady states. These models will include a linear problem, the Brusselator equation, an equation developed by Schnackenberg, and a simple multi-steady state problem. This chapter will conclude by discussing previously developed models used to describe transcription and translation. These models include the model developed by Smolen et al. and incorporated by Hasty, Hasty's model, and Chens larger system model. By better understanding these models, it will be easier to build upon and reformulate these models so as to construct models that represent Air Force needs.

## *3.2 Analysis*

Mathematical models of genetic processes can evolve into systems of ordinary differential equations. When evaluated, these systems can provide detail as to how the model behaves. Techniques for analyzing systems of ordinary differential equations have been available for many years. These techniques involve finding the steady state solution and checking for stability. The principal reference used for this analysis is Brauer and Nohel [5]. They lay the foundation for the analysis of

systems of autonomous ordinary differential equations by giving a general definition of stability, defining stability of linear systems, and explaining how to analyze the stability of nonlinear systems through the process of linearization. Gray and Scott [11] also analyze stability. They give a thorough analysis of how to linearize a complex nonlinear system to analyze stability inside a small region. This analysis is then applied to the metabolism by Heinrich, Rapoport, and Rapoport, as an example of how important this analysis is to the understanding of biological systems.

Brauer and Nohel [5] give the definition of both stability, asymptotic stability, and instability. They begin by defining the autonomous system below.

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}) \text{ where } \mathbf{y} \text{ is a } n \text{ dimensional vector} \tag{3.1}$$

They then define the steady state point which they refer to as the critical point.

$$\text{Let } \mathbf{y} = \mathbf{y}_0 \text{ be a critical point. Then } \mathbf{f}(\mathbf{y}_0) = 0 \tag{3.2}$$

If there exist no other critical points within a neighborhood of $\mathbf{y}_0$, then $\mathbf{y}_0$ is an isolated critical point.

Next, they outline the definition of stability followed by the definition of asymptotic stability.

**Definition 1** *The steady state solution* $\mathbf{y}_0$ *of equation (3.1) is said to be stable if for each number* $\varepsilon > 0$ *we can find a number* $\delta(\varepsilon) > 0$ *such that if* $\boldsymbol{\psi}(t)$ *is any solution of equation (3.1) having* $\|\boldsymbol{\psi}(t_0) - \mathbf{y}_0\| < \delta$*, then the solution* $\boldsymbol{\psi}(t)$ *exists for all* $t \geq t_0$*.*

**Definition 2** *The steady state solution* $\mathbf{y}_0$ *is said to be asymptotically stable if it is stable and if there exists a number* $\delta(\varepsilon) > 0$ *such that if* $\boldsymbol{\psi}(t)$ *is any solution of equation (3.1) having* $\|\boldsymbol{\psi}(t_0) - \mathbf{y}_0\| < \delta$*, then* $\lim_{t\to\infty} \boldsymbol{\psi}(t) = \mathbf{y}_0$*.*

**Definition 3** *The steady state solution* $\mathbf{y}_0$ *is said to be unstable if it is not stable.*

Brauer and Nohel then explain that the simplest general systems for which stability is completely defined is the linear system. Given the linear system in equation (3.3) below, were $A$ is a real constant $n \times n$ matrix, stability can be determined by considering the eigenvalues of $A$.

$$\mathbf{y}' = A\mathbf{y} \tag{3.3}$$

The specific criteria for stability about the origin is given in the following theorem.

**Theorem 1** *"If all eigenvalues of A have nonpositive real parts and all those eigenvalues with zero real parts are simple, then the solution* $\mathbf{y} = 0$ *of (3.3) is stable. If and only if all eigenvalues of A have negative real parts, the zero solution of (3.3) is asymptotically stable. In fact, in this case if* $\boldsymbol{\psi}(t, t_0)$ *denotes the fundamental matrix of (3.3) which is the identity at* $t = t_0$*, then* $\boldsymbol{\psi}(t, t_0) = exp((t - t_0)A)$ *and there exist constants* $K > 0$, $\sigma > 0$ *such that* $|\boldsymbol{\psi}(t, t_0)| \leq K \exp(-\sigma(t - t_0))$ *where* $(t_0 \leq t < \infty)$ *with* $\sigma > 0$ *in the case that all eigenvalues of A have negative real parts and* $\sigma = 0$ *if there are simple eigenvalues with zero real part. If one or more eigenvalues of A have a positive real part, the zero solution of (3.3) is unstable." [5:p 151]*

This analysis of linear systems is useful in describing localized steady states stability of some nonlinear systems. This is done though the process of linearization. Gray and Scott [11] followed the definition provided by Brauer and Nohel [5] by explaining that local asymptotic stability is defined by the behavior of a system very near the steady state point. If a small perturbation from steady state grows, the steady state is unstable. Where as if it decays, the system is stable. The qualification 'local' implies the system may exhibit different stability behavior for larger perturbations. They then explain that linearization is a perturbation method that allows nonlinear systems to be treated locally as linear systems. Brauer and Nohel [5] give a general description of how to linearize a two dimensional system

about a steady state point located at the origin. Gray and Scott give a more explicit derivation of how to linearize the two dimensional system about any steady state point. This description is expanded upon by Heinrich, Rapoport, and Rapoport [15] into an $n$ dimensional case. However, Heinrich, Rapaport, and Rapaport give little detail leading to their results. The detail of the $n$ dimensional case, is given below.

First, consider the nonlinear system given by equation (3.1). Assume the system has a steady state value at $\mathbf{f}(\mathbf{y}_{ss}) = \mathbf{0}$. Now, a small perturbation away from the steady state is given below.

$$\mathbf{y} = \mathbf{y}_{ss} + \Delta\mathbf{y} \text{ where } |\Delta y_i| << 1 \text{ for } (i = 1...n) \tag{3.4}$$

Then the rate of change of the perturbation $\Delta\mathbf{y}$ is obtained by substituting equation (3.4) into equation (3.1). Note that $\mathbf{y}_{ss}$ is a constant so $\frac{d(\mathbf{y}_{ss})}{dt} = 0$.

$$\frac{d(\mathbf{y}_{ss} + \Delta\mathbf{y})}{dt} = \frac{d(\mathbf{y}_{ss})}{dt} + \frac{d(\mathbf{\Delta y})}{dt} = \frac{d(\mathbf{\Delta y})}{dt} = \mathbf{f}(\mathbf{y}_{ss} + \Delta\mathbf{y}) \tag{3.5}$$

Using a Taylor Series Expansion about the steady state point, this equation expands to the following.

$$\frac{d(\Delta y_j)}{dt} = f_j(\mathbf{y}_{ss}) + \sum_{i=1}^{n} \frac{\partial f_j(\mathbf{y}_{ss})}{\partial y_i} \Delta y_j + \text{higher order terms} \quad \text{for } (j = 1...n) \tag{3.6}$$

Now $f_j(\mathbf{y}_{ss}) = \mathbf{0}$ for all $j$ and (higher order terms $\cong 0$) by perturbation theory coupled with the assumption $|\Delta y_i| << 1$ for $(i = 1...n)$ so equation (3.6) simplifies to the following.

$$\frac{d(\Delta y_j)}{dt} = \sum_{i=1}^{n} \frac{\partial f_j(\mathbf{y}_{ss})}{\partial y_i} \Delta y_j \quad \text{for } (j = 1...n) \tag{3.7}$$

Now by introducing the Jacobian matrix $J$, the above equation is rewritten in vector form below.

$$\frac{d(\mathbf{\Delta y})}{dt} = J(\mathbf{y}_{ss})\Delta\mathbf{y} \quad \text{for } (j = 1...n) \tag{3.8}$$

$$\text{where } J(\mathbf{y}) = \begin{pmatrix} \frac{\partial f_1(\mathbf{y})}{\partial y_1} & \frac{\partial f_1(\mathbf{y})}{\partial y_2} & \dots & \frac{\partial f_1(\mathbf{y})}{\partial y_n} \\ \frac{\partial f_2(\mathbf{y})}{\partial y_1} & \frac{\partial f_2(\mathbf{y})}{\partial y_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \frac{\partial f_n(\mathbf{y})}{\partial y_1} & \dots & \dots & \frac{\partial f_n(\mathbf{y})}{\partial y_n} \end{pmatrix} \tag{3.9}$$

Furthermore, $J(\mathbf{y}_{ss})$ is a constant matrix, so equation (3.8) is similar to equation (3.3), and theorem 1 applies. The eigenvalues of the system above are denoted as $\lambda$ in the system below.

$$Det\,|\lambda I - J(\mathbf{y}_{ss})| = 0 \quad \text{where } I \text{ is an } n \times n \text{ Identity matrix} \tag{3.10}$$

Solving for $\lambda$ and applying theorem 1 defines the local steady state stability for the nonlinear system.

Gray and Scott [11] focused on the linearization of a two dimensional model to evaluate various types of stability while retaining a mathematical base. The linearized version of their two dimensional system is as follows.

$$\begin{aligned} \frac{d(\Delta\alpha)}{dt} &= \frac{\partial f_{ss}}{\partial\alpha}\Delta\alpha + \frac{\partial f_{ss}}{\partial\beta}\Delta\beta = a_{11}\Delta\alpha + a_{12}\Delta\beta \\ \frac{d(\Delta\beta)}{dt} &= \frac{\partial g_{ss}}{\partial\alpha}\Delta\alpha + \frac{\partial g_{ss}}{\partial\beta}\Delta\beta = a_{21}\Delta\alpha + a_{22}\Delta\beta \end{aligned} \tag{3.11}$$

Then, using equation (3.10) to solve for the eigenvalues, the quadratic equation below is formed

$$\lambda^2 - tr(J)\lambda + \det(J) = 0 \tag{3.12}$$

where $J = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $tr(J) = a_{11} + a_{22}$, and $\det(J) = a_{11}a_{22} - a_{12}a_{21}$.

Now, equation (3.12) has the solutions $\lambda_{1,2} = \frac{1}{2}\{tr(J) \pm \sqrt{tr(J)^2 - 4\det(J)}\}$. The different possible eigenvalue solutions provide for a number of different types of stability scenarios. These scenarios along with the requirements for their existence are listed in the book *Chemical Oscillations and Instabilities* by Gray and Scott [11:pp 65-68].

### *3.3 Specific cases of stability*

In this section simple mathematical models that exhibit different characteristics of stability are considered. The first model is a simple two component system that displays basic characteristics of stability that are sensitive to different parameter values. A sensitivity analysis is performed on a parameter within the model to examine its effect on the models stability. Also considered are two models of reaction equations that exhibit oscillatory steady state behavior. These models include the Brusselator equation, and an equation developed by Schnackenberg. The last model included will consist of a small system that shows interesting behavior by presenting multiple steady states with different parameter values. By looking at some basic mathematical models that define different types of stability and then incorporating them into biological systems to achieve similar stability characteristics, more can be understood about the biological system as a whole.

*3.3.1 A two component system.* Consider the following two component system.

$$\overset{k_3}{\leftarrow} X \overset{\downarrow k_5}{\underset{k_1}{\overset{k_2}{\rightleftharpoons}}} Y \overset{\downarrow k_6}{\overset{k_4}{\rightarrow}} \tag{3.13}$$

This is a small system that can be involved in any larger biological system. It can be represented as a set of differential equations in the following manner.

$$\frac{d[X]}{dt} = k_1[Y] - k_2[X] - k_3[X] + k_5 \tag{3.14}$$
$$\frac{d[Y]}{dt} = k_2[X] - k_1[Y] - k_4[Y] + k_6 \tag{3.15}$$

This set of differential equations can be simplified to the following set of equations.

$$\begin{aligned} \frac{d[X]}{dt} &= k_1[Y] - k_3^*[X] + k_5 \\ \frac{d[Y]}{dt} &= k_2[X] - k_4^*[Y] + k_6 \end{aligned} \tag{3.16}$$

where $k_3^* = k_2 + k_3$ and $k_4^* = k_1 + k_4$

Now, the stability associated with the above equations can change dramatically given small changes to specific rate constants. For example, say the rate constants are defined as follows, where there is a small perturbation that changes the rate constants in a manner defined by $\varepsilon$.

With $k_1 = 3$, $k_2 = \frac{(2+\varepsilon)^2}{4}$, $k_3^* = 3 + \varepsilon$, $k_4^* = 1$, $k_5 = 0$, $k_6 = 0$, $x = [X]$, $y = [Y]$,

$$V = \begin{pmatrix} x \\ y \end{pmatrix}, \text{ and } J = \begin{pmatrix} -(3+\varepsilon) & 3 \\ \frac{(2+\varepsilon)^2}{4} & -1 \end{pmatrix}, \tag{3.17}$$

then system (3.16) becomes

$$\frac{dV}{dt} = JV \tag{3.18}$$

with a steady state point at $x = 0$ and $y = 0$.

The solution for this system takes the general form

$$X = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t} \tag{3.19}$$

$$Y = c_3 e^{\lambda_1 t} + c_4 e^{\lambda_2 t} \tag{3.20}$$

where $\lambda_1$ and $\lambda_2$ are the eigenvalues of equation (3.18), and are determined using equation (3.12) to be $\lambda_{1,2} = \frac{1}{2}\{tr(J) \pm \sqrt{tr(J)^2 - 4\det(J)}\}$, where $tr(J) = -(4+\varepsilon)$, and $\det(J) = (3+\varepsilon) - 3\frac{(2+\varepsilon)^2}{4} = -2\varepsilon - \frac{3}{4}\varepsilon^2$

$$\text{Thus, } \lambda_{1,2} = \frac{1}{2}\{-(4+\varepsilon) \pm \sqrt{4\,(2+\varepsilon)^2}\} = -\frac{1}{2}(4+\varepsilon) \pm (2+\varepsilon) \tag{3.21}$$

$$\text{So } \lambda_1 = \frac{1}{2}\varepsilon, \text{ and } \lambda_2 = -4 - \frac{3}{2}\varepsilon \tag{3.22}$$

In this system, $\varepsilon$ is restricted to perturbations no less than $-3$ because the kinetic constant $k_3^*$ is restricted to positive values. Therefore, within the interval that $\varepsilon$ is restricted to, there exists two bifurcation points. The first is at $\varepsilon = 0$, and the second is at $\varepsilon = -\frac{8}{3}$. At each bifurcation point, the system changes stability. When $\varepsilon$ is within the interval $[-3, -\frac{8}{3})$, $\lambda_1 < 0$, and $\lambda_2 > 0$. So, the steady state point acts as an unstable saddle point. When $\varepsilon$ is within the interval $(-\frac{8}{3}, 0)$, $\lambda_1 < 0$, and $\lambda_2 < 0$. So, the system converges asymptotically to the steady state point. When $\varepsilon$ is within the interval $(0, \infty)$, $\lambda_1 > 0$, and $\lambda_2 < 0$. So again, the steady state point acts as an unstable saddle point. This example shows that a small change in a single parameter can change the stability of a system greatly. In cases where parameters are a bit relaxed, or changed due to an outside influence, a stability analysis is an important tool in describing the system.

*3.3.2 The Brusselator.* Limit cycle behavior can exist in both chemical systems [9:pp 2352-2357] and biological systems [15]. For this reason, it is necessary to consider the modeling techniques associated with limit cycle oscillatory behavior. First proposed by Prigogine and Lefever [27] in 1967, the Brusselator is a popular chemical model [9] that describes limit cycle oscillatory behavior. It is given by the following chemical equations.

$$A \longrightarrow X \tag{3.23}$$

$$B + X \longrightarrow Y + D \tag{3.24}$$

$$2X + Y \longrightarrow 3X \tag{3.25}$$

$$X \longrightarrow E \tag{3.26}$$

These chemical equations can be transformed into a system of differential equations using the techniques summarized in Chapter 2 and presented by Heinrich, Rapoport, and Rapoport [15]. First, consider the following rate equations describing the reaction equations (3.23)-(3.26).

$$v_1 = -[A] \tag{3.27}$$

$$v_2 = -[B][X] \tag{3.28}$$

$$v_3 = -[X]^2[Y] \tag{3.29}$$

$$v_4 = -[X] \tag{3.30}$$

These can be used to describe the rate of change of each component as it relates to the entire model. This results in the following set of differential equations.

$$\frac{d[X]}{dt} = -v_1 + v_2 + 2v_3 - 3v_3 + v_4 \tag{3.31}$$

$$\frac{d[Y]}{dt} = -v_2 + v_3 \tag{3.32}$$

By assuming $[A] = a$, $[B] = b$, $[X] = x$, and $[Y] = y$ where $a$ and $b$ are held constant, this set of equations simplifies to the following.

$$\frac{dx}{dt} = a - (b+1)x + x^2y \tag{3.33}$$

$$\frac{dy}{dt} = bx - x^2y \tag{3.34}$$

The system is in steady state when $\frac{dx}{dt} = 0$ and $\frac{dy}{dt} = 0$. Therefore, solving for the steady state solution involves simply solving the following system for $x$ and $y$.

$$0 = a - (b+1)x + x^2y \tag{3.35}$$

$$0 = bx - x^2y \tag{3.36}$$

From equation (3.36), $y = \frac{b}{x}$. Substituting $y$ into equation (3.35) yields $x = a$, which yields $y = \frac{b}{a}$. Therefore, the steady state point is as follows.

$$x = a \tag{3.37}$$

$$y = \frac{b}{a} \tag{3.38}$$

As a result of linearizing equations (3.33) and (3.34), the local stability around the steady state can be determined. From equation (3.11),

$$J = \begin{pmatrix} 2xy-(b+1) & x^2 \\ b-2xy & -x^2 \end{pmatrix} \Rightarrow J_{ss} = \begin{pmatrix} b-1 & a^2 \\ -b & -a^2 \end{pmatrix}.$$

The eigenvalues are calculated by solving the characteristic equation 3.39.

$$Det\,|\lambda I - J_{ss}| = Det \begin{vmatrix} \lambda - b + 1 & -a^2 \\ b & \lambda + a^2 \end{vmatrix} = \lambda^2 + \lambda(a^2 + 1 - b) + a^2 \qquad (3.39)$$

Thus, the eigenvalues are $\lambda_{1,2} = \frac{1}{2}\{b - 1 - a^2 \pm \sqrt{(a^2+1-b)^2 - 4a^2}\}$. By setting $b = a^2 + 1$,

$$\lambda_{1,2} = \pm\sqrt{-a^2} = \pm ia, \qquad (3.40)$$

and $\lambda_{1,2}$ are both purely imaginary. This is known as a Poincaré-Andronov-Hopf bifurcation point [12] because any perturbation in $b$ such that $b = a^2 + 1 \pm \varepsilon$ for $\varepsilon > 0$ will lead to either stability or divergent instability. This can be tested through stability analysis. If $b = a^2 + 1 - \varepsilon$, then $\lambda_{1,2} = \frac{1}{2}\{-\varepsilon \pm \sqrt{\varepsilon^2 - 4a^2}\}$ which leads to a damped oscillatory approach to steady state due to the negative real part of $\lambda$. If $b = a^2 + 1 + \varepsilon$, then $\lambda_{1,2} = \frac{1}{2}\{\varepsilon \pm \sqrt{\varepsilon^2 - 4a^2}\}$ which leads to oscillatory divergence due to the positive real part of $\lambda$.

*3.3.3 An Example developed by Schnackenberg.* Schnackenberg[30] systems are simple chemical oscillators. Consider the following model presented by Schnackenberg.

$$2X + Y \xrightarrow{k_1} 3X \tag{3.41}$$

$$A \xrightarrow{k_2} Y \tag{3.42}$$

$$X \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} B \tag{3.43}$$

The velocity though each chemical equation is given as,

$$\begin{aligned} v_1 &= -k_1[X]^2[Y] \\ v_2 &= -k_2[A] \\ v_3 &= -k_3[X] + k_{-3}[B] \end{aligned} \tag{3.44}$$

The differential equations corresponding to (3.41)-(3.43) are

$$\begin{aligned} \frac{d[X]}{dt} &= 2v_1 - 3v_1 + v_3 = -v_1 + v_3 \\ \frac{d[Y]}{dt} &= v_1 - v_2 \end{aligned} \tag{3.45}$$

Using equation (3.44) and letting $[X] = x$, $[Y] = y$, $[A] = \alpha$, and $[B] = \beta$ yields the following.

$$\begin{aligned} \frac{dx}{dt} &= k_1 x^2 y - k_3 x + k_{-3}\beta \\ \frac{dy}{dt} &= -k_1 x^2 y + k_2 \alpha \end{aligned} \tag{3.46}$$

Nondimensionalizing equation (3.46) by setting $x = X_0\xi$, $y = Y_0\eta$, and $t = T_0\tau$ yields the following.

$$\frac{d\xi}{d\tau} = k_1 X_0 Y_0 T_0^2 \xi^2 \eta - k_3 T_0 \xi + \frac{k_{-3} T_0 \beta}{X_0} \tag{3.47}$$

$$\frac{d\eta}{d\tau} = -k_1 X_0^2 T_0 \xi^2 \eta + \frac{k_2 T_0 \alpha}{Y_0} \tag{3.48}$$

If $T_0 = \frac{1}{k_3}$, $X_0 = \sqrt{\frac{k_3}{k_1}}$, and $Y_0 = X_0$, then $k_1 X_0 Y_0 T_0 = 1$, $k_1 X_0^2 T_0 = 1$, $k_3 T_0 = 1$ and equations (3.47) and (3.48) become

$$\frac{d\xi}{d\tau} = \xi^2 \eta - \xi + b \tag{3.49}$$

$$\frac{d\eta}{d\tau} = -\xi^2 \eta + a \tag{3.50}$$

where $b = \frac{k_{-3}}{k_3}\sqrt{\frac{k_1}{k_3}}\beta$ and $a = \frac{k_2}{k_3}\sqrt{\frac{k_1}{k_3}}\alpha$.

The steady state can be found by setting $\frac{d\xi}{dt} = 0$, and $\frac{d\eta}{dt} = 0$, and then solving the following system for $\xi$ and $\eta$ in terms of $a$ and $b$.

$$0 = \xi^2 \eta - \xi + b \tag{3.51}$$

$$0 = -\xi^2 \eta + a \tag{3.52}$$

By adding equation (3.52) to equation (3.51), the equation $0 = -\xi + b + a$, can be solved in terms of $\xi$ to obtain $\xi = a + b$. Then substituting $\xi$ back into equation (3.52), yields $0 = -(a+b)^2\eta + a$, which yields $\eta = \frac{a}{(a+b)^2}$. Thus, the steady state solution is

$$\xi = a + b \tag{3.53}$$

$$\eta = \frac{a}{(a+b)^2} \tag{3.54}$$

Then, linearizing equations (3.49) and (3.50), using equations (3.11) yields,

$$J = \begin{pmatrix} 2\xi\eta - 1 & \xi^2 \\ -2\xi\eta & -\xi^2 \end{pmatrix} \Longrightarrow J_{ss} = \begin{pmatrix} \frac{2a}{(a+b)} - 1 & (a+b)^2 \\ \frac{-2a}{(a+b)} & -(a+b)^2 \end{pmatrix} \tag{3.55}$$

Solving the characteristic equation,

$$\begin{aligned} Det\,|\lambda I - J_{ss}| &= Det \begin{vmatrix} \lambda - \frac{2a}{(a+b)} + 1 & -(a+b)^2 \\ \frac{2a}{(a+b)} & \lambda + (a+b)^2 \end{vmatrix} \\ &= \lambda^2 + \lambda((a+b)^2 + 1 - \frac{2a}{(a+b)}) + (a+b)^2 = 0 \end{aligned} \tag{3.56}$$

yields the eigenvalues $\lambda_{1,2} = \frac{1}{2}\{-(a+b)^2 - 1 + \frac{2a}{(a+b)} \pm \sqrt{((a+b)^2 + 1 - \frac{2a}{(a+b)})^2 - 4(a+b)^2}\}$

By setting $(a+b)^2 + 1 - \frac{2a}{(a+b)} = 0$, the system has a Poincaré-Andronov-Hopf bifurcation point where $(a+b)^3 = a - b$ because the real part of the eigenvalues disappears thereby leaving purely imaginary eigenvalues and therefore, a limit cycle. For a small perturbation from the bifurcation point in one direction, the steady state acts as a spiral sink due to a small negative real part added to the eigenvalues, while for a perturbation in the other direction, the bifurcation point acts as a spiral source due to a small positive real part added to the eigenvalues.

In evaluating the simple Schnackenberg model, there exist many similarities to the Brusselator model. Both models include an autocatalytic chemical equation,

both models measure the rate of change of two species in very small systems, and both models have limit cycles. If the Brusselator were mathematically equivalent to Schnackenbergs model, then there would be a one-to-one onto mapping from the Brusselator model to Schnackenbergs model defining an isomorphism. However, the models are distinctly different. Schnackenberg states that his oscillator is "not isomorphic to the Brusselator and chemically originates from a different and smaller system of reactions." [30]

*3.3.4 A small system involving bistability.* Under specific conditions, a system can have multiple steady states. This sometimes translates to multiple stable steady states. The next system "shows multiple steady states and has been proposed by Schlögl (1972) as a particularly simple and analytically solvable example." [30] It is interesting because its simplicity allows it to be involved in larger systems. The system is as follows.

$$2X + A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} 3X \tag{3.57}$$

$$X \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} B \tag{3.58}$$

The velocity of a single component through these chemical equations is given below

$$\begin{aligned} v_1 &= -k_1[X]^2[A] + k_{-1}[X]^3 \\ v_2 &= -k_2[X] + k_{-2}[B] \end{aligned} \tag{3.59}$$

This allows for each component to be tracked through the model allowing a system of differential equations to be created.

$$\frac{d[X]}{dt} = 2v_1 - 3v_1 + v_2 = -v_1 + v_2 \tag{3.60}$$

Now, substituting equation (3.59) into equation (3.60), the rate of change of the one component system can be tracked explicitly.

$$\frac{d[X]}{dt} = k_1[X]^2[A] - k_{-1}[X]^3 - k_2[X] + k_{-2}[B] \tag{3.61}$$

To nondimensionalize, let $[X] = X_0 x$, $t = T_0\tau$, $[A] = A$, and $[B] = B$ so that

$$\frac{dx}{d\tau} = k_1 X_0 T_0 [A] x^2 - k_{-1} X_0^2 T_0 x^3 - k_2 T_0 x + \frac{k_{-2} T_0 [B]}{X_0} \tag{3.62}$$

If $T_0 = \frac{1}{k_2}$ and $X_0 = \sqrt{\frac{k_2}{k_{-1}}}$, then $k_{-1} X_0^2 T_0 = 1$, $k_2 T_0 = 1$ and

$$\frac{dx}{d\tau} = ax^2 - x^3 - x + b \tag{3.63}$$

where $a = k_1 X_0 T_0 [A]$, and $b = \frac{k_{-2}}{k_2}\sqrt{\frac{k_{-1}}{k_2}}[B]$

Solving for steady state, leads to a cubic polynomial with three roots.

$$x^3 - ax^2 + x - b = 0 \tag{3.64}$$

Understanding oscillatory behavior and multiple steady state behavior in mathematical systems may lead to an understanding of how to detect and construct the behavior in biological models. With this knowledge, it may be possible to gain more understanding of a large biological systems by incorporating these models that represent small systems into them.

*3.4 Smolen et al. and Hasty et al.*

Smolen, Baxter, and Byrne[32] developed a model to test the dynamic properties of a genetic regulatory system. They proposed a model that "captures the salient features of transcription factors, dimerization, binding, and phosphorylation-dependent regulation of transcription." They evaluated the system to find the steady state concentration of the transcription factor. They then plotted the steady state concentration of the Transcription Factor with respect to a parameter $\alpha$ to show how changes in $\alpha$ generate bistability in the system. Unfortunately, Smolen, Baxter, and Byrne neglect to give any sort of derivation to show how their model leads to their conclusions. Hasty et al.[14] picks up on the conclusions of Smolen, Baxter, and Byrne by using the model to evaluate additive noise. Hasty et al. begins by symbolically writing out the model in the form of a system of reaction equations. However, from there, Hasty et al. jump directly to form their conclusions without including any derivation in their model. It is important to understand the derivation of the model in order to judge its validity. Furthermore, understanding how to construct bistability in this model will help to formulate models that represent future Air Force needs.

The construction of this model begins with Hasty et al.'s system of reaction equations.

$$2X \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} X_2 \tag{3.65}$$

$$D + X_2 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} DX_2 \tag{3.66}$$

$$DX_2 + P \overset{k_3}{\rightarrow} DX_2 + nX \tag{3.67}$$

$$X \overset{k_4}{\rightarrow} A \tag{3.68}$$

This system consists of two fast reactions, (3.65) and (3.66), that represent the binding of two $X$ proteins to form a dimer molecule that then binds to a DNA promoter site to form a DNA promoter complex. Then, with the introduction of RNA polymerase, the third equation, represents both transcription and translation that takes place at rate $kf$ to create n $X$ proteins. Finally, the fourth equation represents degradation of protein $X$ at the slow rate of $kd$. Together, these equations represent a simple model of protein regulation through gene expression. An additional equation not listed by Hasty et al. must also be included to reproduce the results presented in their paper. Equation (3.69) represents "the basal rate of production of protein $X$, i.e., the low baseline expression rate in the absence of a transcription factor." [14] It will be determined if this basal rate of production is necessary for the bistability in Chapter 4.

$$INPUT \xrightarrow{r} X \tag{3.69}$$

These equations can be transformed into a system of differential equations using the techniques outlined by Heinrich, Rapoport, and Rapoport [15] and displayed in chapter 2. In the equations below, $v_i$ represents the flux of a single substrate molecule through one of the chemical equation described above. For example, in equation (3.70), $v_1$ represents the reversible flow of one molecule from substrates to products at a rate of $k_1$ and back again with a rate of $k_{-1}$.

$$v_1 = k_{-1}[X_2] - k_1[X]^2 \tag{3.70}$$

$$v_2 = k_{-2}[DX_2] - k_2[D][X_2] \tag{3.71}$$

$$v_3 = -k_3[DX_2][P] \tag{3.72}$$

$$v_4 = -k_4[X] \tag{3.73}$$

$$v_5 = r \tag{3.74}$$

These equations can then be used to track the flow of each molecule through the whole system. Again, this was described by Heinrich, Rapoport, and Rapoport [15] in Chapter 2. Together, the production, reaction, and degradation of each molecule is tracked through the reaction equations using equations (3.70) through (3.74) and described with the differential equations presented next. For example, two $X$ substrate molecules can be tracked through equation (3.65). They are then lost through reaction, but recreated $n$ times over via equation (3.67), and recycled though the system until they degrade in equation (3.68). At the same time, $X$ molecules are being created at a basal rate described by equation (3.69).

$$\frac{d[X]}{dt} = 2v_1 - nv_3 + v_4 + v_5 \tag{3.75}$$

$$\frac{d[X_2]}{dt} = -v_1 + v_2 \tag{3.76}$$

$$\frac{d[D]}{dt} = v_2 \tag{3.77}$$

$$\frac{d[DX_2]}{dt} = -v_2 + v_3 - v_3 \tag{3.78}$$

The degraded state of protein $X$ that takes the form $A$, is not an active part of the system, and therefore, not considered as a rate that drives the system. Furthermore, the concentration of RNA polymerase $P$ is considered constant over time, $[P] = p_0$, so its rate of change is not included in the system above. Equations (3.75) through (3.78) are then expanded to yield the rate equations.

$$\frac{d[X]}{dt} = 2(k_{-1}[X_2] - k_1[X]^2) + nk_3[DX_2]p_0 - k_4[X] + r \tag{3.79}$$

$$\frac{d[X_2]}{dt} = -(k_{-1}[X_2] - k_1[X]^2) + k_{-2}[DX_2] - k_2[D][X_2] \tag{3.80}$$

$$\frac{d[D]}{dt} = k_{-2}[DX_2] - k_2[D][X_2] \tag{3.81}$$

$$\frac{d[DX_2]}{dt} = -(k_{-2}[DX_2] - k_2[D][X_2]) \tag{3.82}$$

Now, (3.81) and (3.82) are negatives of each other. This leads to

$$\frac{d[D]}{dt} + \frac{d[DX_2]}{dt} = \frac{d([D] + [DX_2])}{dt} = 0 \tag{3.83}$$

or more appropriately

$$[D] + [DX_2] = d_T \text{ where } d_T \text{ is a constant} \tag{3.84}$$

Equation (3.84) says the amount of free DNA promoter sites, $[D]$, and the amount of DNA promoter complexes, $[DX_2]$, makes up the total number of DNA promoter sites, $d_T$, in the system.

At this point, Hasty et al.'s [14] system of reaction equations has been transformed into a system of differential equations that represent the model created by Smolen, Baxter, and Byrne[32]. From here, using appropriate assumptions about the reaction speeds of the chemical equations, (3.65) though (3.68), Hasty et al. were able to simplify the equations, (3.79) through (3.84), down to only one equation who's rate was determined only by parameter constants and concentrations of protein $X$. The assumption was that equations (3.65) and (3.66) "are orders of magnitude faster than transcription and degradation." [14] The assumption that equations (3.65) and (3.66) are fast reactions implies that the contents whose concentrations change only within the confines of those equations can be treated as steady

state. This assumption is important because the concentrations of both $X_2$ and $D$ are represented only in equations (3.65) and (3.66). Furthermore, although the concentration of $DX_2$ is represented in the slow equation (3.67), its concentration remains constant through that equation and only fluctuates with equation (3.66). Thus, $\frac{d[X_2]}{dt} = 0$, $\frac{d[D]}{dt} = 0$, and $\frac{d[DX_2]}{dt} = 0$ because the concentrations of $X_2$, $D$, and $DX_2$ are all assumed to exhibit steady state. With this new information, equations (3.79) through (3.84) become

$$\begin{aligned}
\frac{d[X]}{dt} &= 2(k_{-1}[X_2] - k_1[X]^2) + nk_3[DX_2]p_0 - k_4[X] + r && (3.85)\\
0 &= -(k_{-1}[X_2] - k_1[X]^2) + k_{-2}[DX_2] - k_2[D][X_2] && (3.86)\\
0 &= k_{-2}[DX_2] - k_2[D][X_2] && (3.87)\\
d_T &= [D] + [DX_2] && (3.88)
\end{aligned}$$

This system can be simplified in the following manner. Solving equation (3.87) for $DX_2$ in terms of $D$ and $X_2$ yields

$$[DX_2] = \frac{k_2}{k_{-2}}[D][X_2] \tag{3.89}$$

Then substituting equation (3.89) into equation (3.86) results in

$$0 = -(k_{-1}[X_2] - k_1[X]^2) + k_{-2}\frac{k_2}{k_{-2}}[D][X_2] - k_2[D][X_2] \tag{3.90}$$

Solving equation (3.90) in terms of $X_2$ yields

$$[X_2] = \frac{k_1}{k_{-1}}[X]^2 \tag{3.91}$$

Substituting equations (3.89) and (3.91) into equation (3.85) and simplifying results in

$$\frac{d[X]}{dt} = np_0 \frac{k_1}{k_{-1}} \frac{k_2}{k_{-2}} k_3[D][X]^2 - k_4[X] + r \tag{3.92}$$

Equation (3.88) can then be used to remove the dependence on $[D]$. First substitute equation (3.89) followed by equation (3.91) into equation (3.88) to obtain

$$d_T = [D] + \frac{k_1}{k_{-1}} \frac{k_2}{k_{-2}} [D][X]^2 \tag{3.93}$$

Solving for $[D]$ yields

$$[D] = \frac{d_T}{1 + \frac{k_1}{k_{-1}} \frac{k_2}{k_{-2}} [X]^2} \tag{3.94}$$

This equation can then be inserted into equation (3.92) to eliminate the dependence on $[D]$.

$$\frac{d[X]}{dt} = \frac{np_0 \frac{k_1}{k_{-1}} \frac{k_2}{k_{-2}} k_3 d_T [X]^2}{1 + \frac{k_1}{k_{-1}} \frac{k_2}{k_{-2}} [X]^2} - k_4[X] + r \tag{3.95}$$

Finally, by defining $K_1 = \frac{k_1}{k_{-1}}$, $K_2 = \frac{k_2}{k_{-2}}$, $\alpha = np_0 K_1 K_2 k_3 d_T$, $\beta = K_1 K_2$, $\gamma = k_4$, $\delta = r$, and $u = [X]$, the rate equation for the protein $X$ can be displayed in the form which Hasty et al.[14] and Smolen et al. [32] present it.

$$\frac{du}{dt} = \frac{\alpha u^2}{1 + \beta u^2} - \gamma u + \delta \tag{3.96}$$

Hasty et al., and Smolen et al. then examined how the steady state value of $u$ changed as a function of $\alpha$. They incorporated a bifurcation plot to track the steady state concentration of $u$ with respect to $\alpha$, and discovered that as $\alpha$ increases over a particular interval, $u$ attains multiple steady states and in fact becomes bistable. Again, neither Hasty et al. or Smolen et al. give explicit derivation showing how they

achieved their results. Therefore, we will now analyze steady state concentration of $u$ with respect to $\alpha$ in an attempt to recreate their results.

First, set $\frac{du}{dt} = 0$ in order to evaluate the steady state concentration of $u$.

$$0 = \frac{\alpha u^2}{1 + \beta u^2} - \gamma u + \delta \tag{3.97}$$

Then, multiply through by the denominator and simplify to obtain the following cubic polynomial.

$$0 = u^3 - \frac{(\delta\beta + \alpha)}{\gamma\beta} u^2 + \frac{\gamma}{\gamma\beta} u - \frac{\delta}{\gamma\beta} \tag{3.98}$$

Then, for different parameter settings, this equation can be solved for $u$. For example, Hasty set $\beta = 1$, $\gamma = 10$, and $\delta = .4$ and then solved for $u$ with respect to $\alpha$ over the interval $\alpha = [0, 70]$. With this choice of parameters, equation (3.98) becomes

$$0 = u^3 - \frac{(.4 + \alpha)}{10} u^2 + u - .04 \tag{3.99}$$

For each value of $\alpha$, the roots of this polynomial can be determined.

### *3.5 Hasty et al.*

The Air Force is searching for rapid ways to construct models of cellular systems. They are interested in understanding the unique genetic responses that come from novel Air Force chemicals introduced into cellular systems. One way to quickly create a model that incorporates desired responses is by considering other related models. Hasty et al.[13] shows that it is possible to take a genetic model that incorporates normal cellular activity and, using it as a template, construct a new model that incorporates an elaborate cellular response to outside stimuli. Understanding how Hasty et al. drew from the previous model developed by Smolen [32] to create a

new model will help the Air Force to quickly generate it's own models using similar techniques.

Hasty drew from the model constructed in the previous section to create a model representing two different tracks the bacteriophage $\lambda$ virus may take to infect a cell. Hasty et al. began by writing out the reaction equations that represent the cellular processes taking place. One of the assumptions in [13] is that equations (3.101) though (3.104) represent fast reactions, where as the others represent slow reactions.

$$Input \xrightarrow{r} X \tag{3.100}$$

$$2X \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} X_2 \tag{3.101}$$

$$D + X_2 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} DX_2 \tag{3.102}$$

$$D + X_2 \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} DX_2^* \tag{3.103}$$

$$DX_2 + X_2 \underset{k_{-4}}{\overset{k_4}{\rightleftharpoons}} DX_2X_2 \tag{3.104}$$

$$DX_2 + P \xrightarrow{kf} DX_2 + P + nX \tag{3.105}$$

$$X \xrightarrow{kd} A \tag{3.106}$$

Equation (3.100) is not specifically included in Hasty et al.'s system of chemical equations, but does represent the basal rate of production of $X$ as assumed by Hasty. Note that the previous model is incorporated into this model as equations (3.100), (3.101), (3.102), (3.105), and (3.106). Yet, with the introduction of equations (3.103) and (3.104), this model now can represent Hasty et al's mutant system describing the lysis - lysogeny decision of the bacteriophage $\lambda$ virus.

The analysis of this model is much the same as the previous model. It begins with equations representing the flux of one molecule through each equation.

$$v_1 = r \tag{3.107}$$

$$v_2 = k_{-1}[X_2] - k_1[X]^2 \tag{3.108}$$

$$v_3 = k_{-2}[DX_2] - k_2[D][X_2] \tag{3.109}$$

$$v_4 = k_{-3}[DX_2^*] - k_3[D][X_2] \tag{3.110}$$

$$v_5 = k_{-4}[DX_2X_2] - k_4[DX_2][X_2] \tag{3.111}$$

$$v_6 = -k_T[DX_2][P] \tag{3.112}$$

$$v_7 = -k_d[X] \tag{3.113}$$

These equations can then be incorporated into differential equations that represent the flow of each molecule through the system. The simplified form of this system is given below.

$$\frac{d[X]}{dt} = v_1 + 2v_2 - nv_6 + v_7 \tag{3.114}$$

$$\frac{d[X_2]}{dt} = -v_2 + v_3 + v_4 + v_5 \tag{3.115}$$

$$\frac{d[D]}{dt} = v_3 + v_4 \tag{3.116}$$

$$\frac{d[DX_2]}{dt} = -v_3 + v_4 \tag{3.117}$$

$$\frac{d[DX_2^*]}{dt} = -v_4 \tag{3.118}$$

$$\frac{d[DX_2X_2]}{dt} = -v_5 \tag{3.119}$$

Then, substituting equations (3.107) through (3.113) into equations (3.114) through (3.119), the following differential equations are created that represent the system.

$$\frac{d[X]}{dt} = r + 2(k_{-1}[X_2] - k_1[X]^2) + nk_T[DX_2][P] - k_d[X]$$

$$\frac{d[X_2]}{dt} = -(k_{-1}[X_2] - k_1[X]^2) + k_{-2}[DX_2] - k_2[D][X_2] + k_{-3}[DX_2^*] - k_3[D][X_2] + k_{-4}[DX_2X_2] - k_4[DX_2][X_2]$$

$$\frac{d[D]}{dt} = k_{-2}[DX_2] - k_2[D][X_2] + k_{-3}[DX_2^*] - k_3[D][X_2]$$

$$\frac{d[DX_2]}{dt} = -(k_{-2}[DX_2] - k_2[D][X_2]) + k_{-3}[DX_2^*] - k_3[D][X_2]$$

$$\frac{d[DX_2^*]}{dt} = -(k_{-3}[DX_2^*] - k_3[D][X_2])$$

$$\frac{d[DX_2X_2]}{dt} = -(k_{-4}[DX_2X_2] - k_4[DX_2][X_2])$$

With the assumption that (3.101) though (3.104) represent fast reactions, steady state approximations can be made regarding some of the equations above. Specifically, $\frac{d[X_2]}{dt} = 0$, $\frac{d[D]}{dt} = 0$, $\frac{d[DX_2]}{dt} = 0$, $\frac{d[DX_2^*]}{dt} = 0$, and $\frac{d[DX_2X_2]}{dt} = 0$. With the substitutions $x = [X]$, $y = [X_2]$, $d = [D]$, $u = [DX_2]$, $v = [DX_2^*]$, $z = [DX_2X_2]$, and $p_0 = [P]$ the following equations result.

$$\frac{dx}{dt} = 2(k_{-1}y - k_1^2x) + np_0k_Tu - k_dx + r \tag{3.120}$$

$$0 = -(k_{-1}y - k_1^2x) + k_{-2}u - k_2dy + k_{-3}v - k_3dy + k_{-4}z - k_4uy \tag{3.121}$$

$$0 = k_{-2}u - k_2dy + k_{-3}v - k_3dy \tag{3.122}$$

$$0 = -(k_{-2}u - k_2dy) + k_{-3}v - k_3dy \tag{3.123}$$

$$0 = -(k_{-3}v - k_3dy) \tag{3.124}$$

$$0 = -(k_{-4}z - k_4uy) \tag{3.125}$$

Substituting equation (3.124) into equation (3.123), results in $0 = -(k_{-2}u - k_2dy)$. Then, substituting this result along with equations (3.123) and (3.124) into equation (3.121) yields $0 = -(k_{-1}y - k_1^2x)$ which can be solved for $y$ as a function of $x$

so that $y = \frac{k_1}{k_{-1}}x^2$. Then substituting $y = \frac{k_1}{k_{-1}}x^2$ into $0 = -(k_{-2}u - k_2 dy)$, and solving for $u$ yields the equation $u = \frac{k_1}{k_{-1}}\frac{k_2}{k_{-2}}x^2 d$. Together, substituting $u = \frac{k_1}{k_{-1}}\frac{k_2}{k_{-2}}x^2 d$ and $y = \frac{k_1}{k_{-1}}x^2$ into equation (3.125) yields $z = (\frac{k_1}{k_{-1}})^2\frac{k_2}{k_{-2}}\frac{k_4}{k_{-4}}x^4 d$. Finally, substituting $y = \frac{k_1}{k_{-1}}x^2$ into equation (3.124), results in the equation $v = \frac{k_1}{k_{-1}}\frac{k_3}{k_{-3}}x^2 d$. Thus, equations (3.121)-(3.125) can be written as the following equations where $y$, $u$, $v$, and $z$ are functions with respect to the terms $d$ and $x$.

$$y = \frac{k_1}{k_{-1}}x^2 \tag{3.126}$$

$$u = \frac{k_1}{k_{-1}}\frac{k_2}{k_{-2}}x^2 d \tag{3.127}$$

$$v = \frac{k_1}{k_{-1}}\frac{k_3}{k_{-3}}x^2 d \tag{3.128}$$

$$z = (\frac{k_1}{k_{-1}})^2\frac{k_2}{k_{-2}}\frac{k_4}{k_{-4}}x^4 d \tag{3.129}$$

Substituting these equations into equation (3.120) and simplifying yields the following equation.

$$\frac{dx}{dt} = np_0 k_T \frac{k_1}{k_{-1}}\frac{k_2}{k_{-2}}x^2 d - k_d x + r \tag{3.130}$$

Now, $\frac{d[D]}{dt} = 0$, $\frac{d[DX_2]}{dt} = 0$, $\frac{d[DX_2^*]}{dt} = 0$, and $\frac{d[DX_2X_2]}{dt} = 0$ implies $\frac{d[D]}{dt} + \frac{d[DX_2]}{dt} + \frac{d[DX_2^*]}{dt} + \frac{d[DX_2X_2]}{dt} = 0$ or more explicitly, $\frac{d[D+DX_2+DX_2^*+DX_2X_2]}{dt} = 0$. This means there exists a constant number of binding sites $d_T$ such that,

$$D + DX_2 + DX_2^* + DX_2X_2 = d_T \tag{3.131}$$

$$\text{which implies } d + u + v + z = d_T \tag{3.132}$$

Now, by substituting equations (3.126) though (3.129) into equation (3.132), the following equation emerges.

$$d + \frac{k_1}{k_{-1}}\frac{k_2}{k_{-2}}x^2 d + \frac{k_1}{k_{-1}}\frac{k_3}{k_{-3}}x^2 d + (\frac{k_1}{k_{-1}})^2\frac{k_2}{k_{-2}}\frac{k_4}{k_{-4}}x^4 d = d_T \tag{3.133}$$

Equation (3.133) can then be solved in terms of $d$ which results in

$$d = \frac{d_T}{1 + \frac{k_1}{k_{-1}}\frac{k_2}{k_{-2}}x^2 + \frac{k_1}{k_{-1}}\frac{k_3}{k_{-3}}x^2 + (\frac{k_1}{k_{-1}})^2\frac{k_2}{k_{-2}}\frac{k_4}{k_{-4}}x^4} \tag{3.134}$$

This equation can then be incorporated into (3.130) to give an equation in terms of $x$.

$$\frac{dx}{dt} = \frac{np_0 k_T \frac{k_1}{k_{-1}}\frac{k_2}{k_{-2}} d_T x^2}{1 + \frac{k_1}{k_{-1}}\frac{k_2}{k_{-2}}x^2 + \frac{k_1}{k_{-1}}\frac{k_3}{k_{-3}}x^2 + (\frac{k_1}{k_{-1}})^2\frac{k_2}{k_{-2}}\frac{k_4}{k_{-4}}x^4} - k_d x + r \tag{3.135}$$

This last equation can be simplified by the substitutions $K_1 = \frac{k_1}{k_{-1}}$, $K_2 = \frac{k_2}{k_{-2}}$, $K_3 = \frac{k_3}{k_{-3}}$, $K_4 = \frac{k_4}{k_{-4}}$ and by further substitution of $K_3 = \sigma_1 K_2$, and $K_4 = \sigma_2 K_2$.

$$\frac{dx}{dt} = \frac{np_0 k_T K_1 K_2 d_T x^2}{1 + (1+\sigma_1)K_1 K_2 x^2 + \sigma_2 (K_1)^2 (K_2)^2 x^4} - k_d x + r \tag{3.136}$$

Then using scaling techniques and the substitutions $x = X_0\xi$, and $t = T_0\tau$, this equation can be rewritten as

$$\frac{d\xi}{d\tau} = \frac{np_0 k_T K_1 K_2 d_T T_0 X_0^2 \xi^2}{1 + (1+\sigma_1)K_1 K_2 X_0^2 \xi^2 + \sigma_2 (K_1)^2 (K_2)^2 X_0^4 \xi^4} - k_d T_0 \xi + \frac{T_0 r}{X_0} \tag{3.137}$$

Then making the substitution $X_0 = \frac{1}{\sqrt{K_1 K_2}}$, and $T_0 = \frac{1}{r\sqrt{K_1 K_2}}$, the equation further simplifies.

$$\frac{d\xi}{d\tau} = \frac{\frac{np_0k_Td_T}{r}\xi^2}{1+(1+\sigma_1)\xi^2+\sigma_2\xi^4} - \frac{k_d}{r\sqrt{K_1K_2}}\xi + 1 \tag{3.138}$$

Finally, setting $\alpha = \frac{np_0k_Td_T}{r}$, and $\gamma = \frac{k_d}{r\sqrt{K_1K_2}}$, Hasty et al. equation is finally realized.

$$\frac{d\xi}{d\tau} = \frac{\alpha\xi^2}{1+(1+\sigma_1)\xi^2+\sigma_2\xi^4} - \gamma\xi + 1 \tag{3.139}$$

## *3.6 Two New Models of Transcription and Translation*

*3.6.1 New Model #1.* Hasty et al.'s second model [13] is an example of how previous models have been extended to incorporate different aspects of behavior in cellular systems. The model presented in this section was constructed in an effort to represent transcription and translation in a similar fashion to Hasty et al.'s model[14]. The goal of this model is to track the concentration of varying amounts of two different effector molecules as they relate to transcription and translation. By incorporating multiple binding of one type of molecule into a protein that effects its own creation, an autoregulatory feedback loop is created that may produce interesting behavior. In this model, if $m = 0$ and $b > 0$, then the model is very similar to Hasty's model [14]. Although this is a hypothetical model, it acts as a first step in a transition from previously created models to the construction of models unique to the Air Force.

$$aE_1 + C_p \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} RNAp \tag{3.140}$$

$$bE_2 + C_r \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \check{R} \tag{3.141}$$

$$\check{R} + RNAp \xrightarrow{k_3} \check{R}RNAp \tag{3.142}$$

$$RNAp + DNA \xrightarrow{k_4} RNAp + nE_1 + mE_2 \tag{3.143}$$

$$E_1 \xrightarrow{k_5} \text{degrades} \tag{3.144}$$

$$E_2 \xrightarrow{k_6} \text{degrades} \tag{3.145}$$

Equations (3.140), (3.141), and (3.142) represent fast reactions. In Equation (3.140), $a$ is an integer representing the number of the particular type of effector molecule $E_1$ that bonds with the molecular complex $C_p$ to form RNA polymerase, $RNAp$. In Equation (3.141), $b$ is an integer representing the number of the particular type of effector molecule $E_2$ that bonds with the molecular complex $C_r$ to form an RNA polymerase repressor, $\check{R}$. Equation (3.142), represents the bonding of the repressor $\check{R}$ to the RNA polymerase $RNAp$ to create a complex $\check{R}RNAp$ that prevents the RNA polymerase from binding to the DNA and activating transcription and translation.

Equations (3.143) through (3.145) are slow reactions. In Equation (3.143), the RNA polymerase $RNAp$ attaching to the $DNA$ promoter site to initiate transcription and thereby initiate translation which continues until the RNA degrades. The end products are the $RNAp$, as well as, the two effector molecules: $E_1$, that was created $n$ times in translation, and $E_2$, that was created $m$ times in translation. Equations (3.144) and (3.145) represent the degradation of the proteins $E_1$ and $E_2$.

Certain assumptions about the model are made immediately. First, it is assumed that the number of molecular complexes, $C_p$ and $C_r$, are constant. It is also assumed that the number of $DNA$ promoter sites is constant. With this, the

following equations represent the flux of one molecule through each of the equations (3.140)-(3.145).

$$\begin{aligned} v_1 &= k_{-1}[RNAp] - k_1 C_p [E_1]^a \\ v_2 &= k_{-2}[\check{R}] - k_2 C_r [E_2]^b \\ v_3 &= -k_3 [RNAp][\check{R}] \\ v_4 &= -k_4 DNA[RNAp] \\ v_5 &= -k_5 [E_1] \\ v_6 &= -k_6 [E_2] \end{aligned} \tag{3.146}$$

These equations can be used to formulate a system of differential equations defining the rate of change of molecular concentrations through the system. This system of differential equations is

$$\begin{aligned} \frac{d[E_1]}{dt} &= av_1 - nv_4 + v_5 \\ \frac{d[E_2]}{dt} &= bv_2 - mv_4 + v_6 \\ \frac{d[RNAp]}{dt} &= -v_1 + v_3 + v_4 - v_4 = -v_1 + v_3 \\ \frac{d[\check{R}]}{dt} &= -v_2 + v_3. \end{aligned} \tag{3.147}$$

By substituting the system (3.146) into the system (3.147), a complete system is generated.

$$\begin{aligned}
\frac{d[E_1]}{dt} &= ak_{-1}[RNAp] - ak_1C_p[E_1]^a + nk_4DNA[RNAp] - k_5[E_1] \quad (3.148)\\
\frac{d[E_2]}{dt} &= bk_{-2}[\check{R}] - bk_2C_r[E_2]^b + mk_4DNA[RNAp] - k_6[E_2]\\
\frac{d[RNAp]}{dt} &= -k_{-1}[RNAp] + k_1C_p[E_1]^a - k_3[RNAp][\check{R}]\\
\frac{d[\check{R}]}{dt} &= -k_{-2}[\check{R}] + k_2C_r[E_2]^b - k_3[RNAp][\check{R}]
\end{aligned}$$

Now, by nondimensionalizing, this problem can be simplified. Consider the following substitutions. $t = T_0\tau$, $[E_1] = E_{1_0}e_1$, $[E_2] = E_{2_0}e_2$, $[RNAp] = P_0p$, and $[\check{R}] = R_0r$, where $\frac{d}{dt} = \frac{d}{d\tau}\frac{d\tau}{dt} = \frac{1}{T_0}\frac{d}{d\tau}$.

$$\begin{aligned}
\frac{d[E_1]}{dt} &= a\frac{k_{-1}T_0P_0}{E_{1_0}}p - ak_1C_pT_0E_{1_0}^{a-1}e_1^a + n\frac{k_4DNAT_0P_0}{E_{1_0}}p - k_5T_0e_1 \quad (3.149)\\
\frac{d[E_2]}{dt} &= b\frac{k_{-2}T_0R_0}{E_{2_0}}r - bk_2C_rT_0E_{2_0}^{b-1}e_2^b + m\frac{k_4DNAT_0P_0}{E_{2_0}}p - k_6T_0e_2\\
\frac{d[RNAp]}{dt} &= -k_{-1}T_0p + \frac{k_1C_pT_0E_{1_0}^a}{P_0}e_1^a - k_3T_0R_0pr\\
\frac{d[\check{R}]}{dt} &= -k_{-2}T_0r + \frac{k_2C_rT_0E_{2_0}^b}{R_0}e_2^b - k_3T_0P_0pr
\end{aligned}$$

To simplify system (3.149), set $k_{-2}T_0 = 1$ so that $T_0 = \frac{1}{k_{-2}}$. Then set $k_1C_pT_0E_{1_0}^{a-1} = 1$, which yields $E_{1_0} = \sqrt[a-1]{\frac{k_{-2}}{k_1C_p}} = \xi$. With the requirement that $k_2C_rT_0E_{2_0}^{b-1} = 1$, $E_{2_0} = \sqrt[b-1]{\frac{k_{-2}}{k_2C_r}} = \eta$. Requiring $\frac{k_{-1}T_0P_0}{E_{1_0}} = 1$, yields $P_0 = \frac{k_{-2}}{k_{-1}}\xi$. Finally, requiring $\frac{k_{-2}T_0R_0}{E_{2_0}} = 1$, requires that $R_0 = \sqrt[b-1]{\frac{k_{-2}}{k_2C_r}} = \eta$.

This leads to a simplification of the original system (3.148).

$$\frac{de_1}{dt} = a(p - e_1^a) + n\frac{k_4 DNA}{k_{-1}}p - \frac{k_5}{k_{-2}}e_1 \tag{3.150}$$

$$\frac{de_2}{dt} = b(r - e_2^b) + m\frac{k_4 DNA\xi}{\eta k_{-1}}p - \frac{k_6}{k_{-2}}e_2 \tag{3.151}$$

$$\frac{dp}{dt} = -\frac{k_{-1}}{k_{-2}}p + \frac{k_1 C_p k_{-1}\xi^{a-1}}{k_{-2}^2}e_1^a - \frac{k_3\eta}{k_{-2}}pr \tag{3.152}$$

$$\frac{dr}{dt} = -r + \frac{k_2 C_r \eta^{b-1}}{k_{-2}}e_2^b - \frac{k_3\xi}{k_{-1}}pr \tag{3.153}$$

Now, because equations (3.140), (3.141), and (3.142) represent fast reactions, and $RNAp$ and $\check{R}$ only change concentrations within those reactions, they can be assumed to be at equilibrium concentrations. Therefore, let $\frac{dp}{dt} = 0$, and $\frac{dr}{dt} = 0$, and solve equations (3.152) and (3.153) for $r$ and $p$ in terms of $e_1$ and $e_2$ to get equations that can be plugged into equations (3.150) and (3.151). Therefore, the problem can be formulated in terms of only $\frac{de_1}{dt}$ and $\frac{de_2}{dt}$.

*3.6.2 New Model #2.* A second problem, where only the general derivation is included, is listed here. This problem is much the same as the previous problem, however, it separates transcription and translation into two equations. Transcription is listed as equation (3.154), and is a slow equation. Translation is listed as equation (3.155), and is a relatively fast equation compared with the degradation of mRNA (denoted $RNA_m$) which is listed as equation (3.156) and is a slow reaction. Equation (3.155) is fast because a large concentration of ribosomes are assumed to be in the system.

$$aE_1 + C_p \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} RNAp$$

$$bE_2 + C_r \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \check{R}$$

$$\check{R} + RNAp \xrightarrow{k_3} \check{R}RNAp$$

$$RNAp + DNA \xrightarrow{k_4} RNAp + RNA_m \tag{3.154}$$

$$RNA_m + Ribosomes \xrightarrow{k_5} RNA_m + Ribosomes + E_1 + E_2 \tag{3.155}$$

$$RNA_m \xrightarrow{k_6} \text{degrades} \tag{3.156}$$

$$E_1 \xrightarrow{k_7} \text{degrades}$$

$$E_2 \xrightarrow{k_8} \text{degrades}$$

Again, certain assumptions about the model are made immediately. First, it is assumed that the number of molecular complexes, $C_p$ and $C_r$, are constant. It is also assumed that the number of $DNA$ promoter sites are constant. Furthermore, the *Ribosomes* are considered to be a large constant. With this, the following equations represent the flux of one molecule through each of the previous equations.

$$\begin{aligned}
v_1 &= k_{-1}[RNAp] - k_1 C_p [E_1]^a \qquad (3.157)\\
v_2 &= k_{-2}[\check{R}] - k_2 C_r [E_2]^b \\
v_3 &= -k_3 [RNAp][\check{R}] \\
v_4 &= -k_4 DNA[RNAp] \\
v_5 &= -k_5 Ribosomes[RNA_m] \\
v_6 &= -k_6 [RNA_m] \\
v_7 &= -k_7 [E_1] \\
v_8 &= -k_8 [E_2]
\end{aligned}$$

These equations can be used to formulate a system of differential equations defining the rate of change of molecular concentrations through the system. This system of differential equations is

$$\begin{aligned}
\frac{d[E_1]}{dt} &= av_1 - v_5 + v_7 \\
\frac{d[E_2]}{dt} &= bv_2 - v_5 + v_8 \\
\frac{d[RNAp]}{dt} &= -v_1 + v_3 + v_4 - v_4 = -v_1 + v_3 \\
\frac{d[\check{R}]}{dt} &= -v_2 + v_3 \\
\frac{d[RNA_m]}{dt} &= -v_4 + v_5 - v_5 + v_6 = -v_4 + v_6.
\end{aligned} \tag{3.158}$$

By substituting the system (3.157) into the system (3.158), a complete system is generated.

$$\begin{aligned}
\frac{d[E_1]}{dt} &= ak_{-1}[RNAp] - ak_1C_p[E_1]^a + k_5 Ribosomes[RNA_m] - k_7[E_1] \\
\frac{d[E_2]}{dt} &= bk_{-2}[\check{R}] - bk_2C_r[E_2]^b + k_5 Ribosomes[RNA_m] - k_8[E_2] \\
\frac{d[RNAp]}{dt} &= -k_{-1}[RNAp] + k_1C_p[E_1]^a - k_3[RNAp][\check{R}] \\
\frac{d[\check{R}]}{dt} &= -k_{-2}[\check{R}] + k_2C_r[E_2]^b - k_3[RNAp][\check{R}] \\
\frac{d[RNA_m]}{dt} &= k_4 DNA[RNAp] - k_6[RNA_m]
\end{aligned}$$

Again, due to the autoregulatory feedback of this system, coupled with the degradation, this equation is expected to show interesting behavior. However, due to the complexity of this problem, it will not be simplified, nor derived further here.

## *3.7 Glucose*

The glycolytic pathway is metabolic pathway that starts with glucose and produces pyruvate via ten catalyzed reactions. "Glycolysis may be considered to occur in two stages. Stage I (Reactions 1-5): Glucose is phosphorylated and cleaved to form two molecules of the triose glyceraldehyde-3-phosphate. This requires the expenditure of two ATPs in an 'energy investment' (Reactions 1 and 3). Stage II (Reactions 6-10): The two molecules of glyceraldehyde-3-phosphate are converted to pyruvate with the concomitant generation of four ATPs (Reaction 7 and 10)." [34:p 446] Figure 3.1 outlines the glycolysis pathway. Using both Gepasi and Matlab, a basic model of the glycolytic pathway was created. The model was created by using reaction equations associated with mass action to define a system of differential equations. Both in Gepasi and Matlab, the model is the same so as to compare the Gepasi simulation with the Matlab simulation.

*3.7.1 Gepasi.* Using Gepasi to create the glycolysis pathway was an easy task. Simply type the twenty catalytic reactions that define the system into the 'reactions' box that is listed on the 'model definition' tab. These equations can be written in the form of reaction equations. Then, going to the 'kinetics' box, the reaction equations can be described with any one of a series of different kinetic types. The Glucose model used mass action. In choosing mass action, boxes appear that allow for the kinetic constants to be incorporated into the reaction. Once the reaction equations are complete, they are converted by Gepasi into differential equations that can be solved by the LSODA solver (Livermore Solver of Ordinary Differential Equations). By incorporating initial conditions in the 'metabolites' box that is listed on the 'model definition' tab, the model is set to run. By going to the 'tasks' tab, an end time can be chosen as well as the number of time steps LSODA will use to approximate the system. By checking the first two boxes in the report section of the 'tasks' tab, the output data from Gepasi will include the concentrations of all reactant species at the end time chosen.

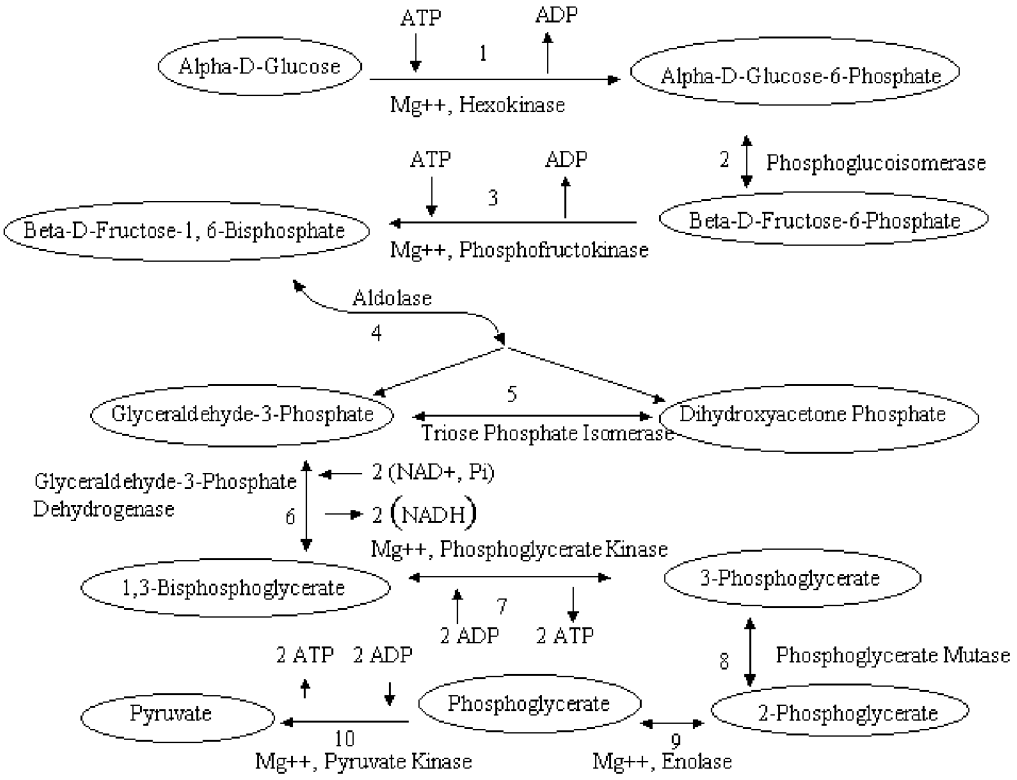


Figure 3.1. The glycolysis metabolic pathway

*3.7.2 MATLAB.* In the Matlab script, there are 36 different molecules that are changing in the system. Each reaction has a reaction rate coefficient associated with it. The reaction rate coefficients are given in the code by $k1-k37$. The forward reaction rate coefficients are 3 and the backwards reaction rate coefficients are 1. In this model, all reaction rate coefficients are fixed. In future models, they may be considered variables or functions. Part of the script is listed below. In reaction #2 there is only one kinetic constant because there is no backward reaction, the reaction is one-way.

$$k1 = 3; \quad \text{REACTION\#1} \tag{3.159}$$

$$k2 = 1; \tag{3.160}$$

$$k3 = 3; \quad \text{REACTION\#2} \tag{3.161}$$

$$k4 = 3; \quad \text{REACTION\#3} \tag{3.162}$$

$$k5 = 1; \tag{3.163}$$

Each reaction is dependent on the molecules formed by previous reactions. Therefore, all reaction in this system are interconnected through time. This interconnection makes modeling the system as a whole, a very difficult process. However, first looking at each reaction individually, then incorporating the equations for the individual reactions into the complete model makes this model easier to construct. There are 10 individual reactions in the very simple biochemical model of the metabolic pathway, Glycolysis. However, each reaction is a catalytic reaction. This being the case, each reaction can be split into two reactions that more accurately describe the catalytic reaction. [18] This changes the 10 reactions into 20 slightly more complex reactions involving hypothetical molecular states of catalytic transition. Each of the 20 reactions has a reaction rate depending on the concentration of molecules entering into the system and the concentration of those leaving. Before

looking at the system as a whole, the individual reaction equations are constructed. Below is a summary of all the reactants of the system. Also included are the first two of twenty reaction equation along with the biochemical equations that explain them.

_____ QUICK KEY _____

ENZYMES = $y(1) - y(9)$, $y(35)$

CATALYST TRANSITION = $y(10) - y(19)$

ENERGY = $y(20) - Y(24)$

SUBSTRATES = $y(25) - y(34)$, $y(36)$

_____

The first two reaction equations in the system explain the reaction process that takes alpha-D-Glucose to alpha-D-Glucose-6-phosphate via the catalyst Hexokinase and coenzyme $Mg^{++}$. In the process, this reaction requires energy in the form of ATP and creates the by-product ADP.

The first reaction equation takes alpha-D-Glucose in the presence of ATP, Hexokinase, and $Mg^{++}$ and produces Catalyst_transition_1. It is presented as follows.

alpha-D-Glucose + ATP + $Mg^{++}$+ Hexokinase $(k2)$ <==> $(k1)$ Catalyst_transition_1

The algebraic equation corresponding to it is

$$eq1 = -k1 * y(25) * y(20) * y(1) * y(2) + k2 * y(10).$$

The second reaction equation describes Catalyst_transition_1 producing alpha-D-Glucose-6-phosphate, ADP, Hexokinase, and $Mg^{++}$. It is presented as follows.

Catalyst_transition_1 => ($k3$) alpha-D-Glucose-6-phosphate + ADP + $Mg^{++}$+ Hexokinase

The algebraic equation corresponding to it is

$$eq2 = -k3 * y(10).$$

The concentration of each of the 36 molecules change with respect to time. Some of these molecules exist in more than one equation. Each of the following equations is a reaction rate written with respect to a single molecule on the left hand side of the biochemical equation. The negative of that equation will give the reaction rate written with respect to a single molecule on the right hand side of the biochemical equation. To find the reaction rate of a specific molecule over the entire system, add each equation where the molecule appears on the left hand side of the biochemical equation. Of course, if the molecule appears on the right hand side of the biochemical equation, add the negative of the above equations.

$$\begin{aligned} \frac{dy(1)}{dt} &= eq1 - eq2 + eq5 - eq6 + eq13 - eq14 + eq17 - eq18 + eq19 - eq20; \\ \frac{dy(2)}{dt} &= eq1 - eq2; \end{aligned} \tag{3.164}$$

The first equation (3.164) from the larger system exhibits a change of Mg++ with respect to time. The second equation (3.164) from exhibits a change of Hexokinase with respect to time. Looking through this system, you will notice that these equations are nonlinear. This makes solving the system exactly, virtually impossible. However, a numerical solver such as Matlab's 'ode23' or 'ode23s' can solve the system easily. The script written for Glucose uses the 'ode23s' stiff solver to solve the system. This stiff solver requires a set of initial conditions for all thirty six reactant species as well as a time interval in order to solve the system. Comparing

the output of the Matlab script at a specific end time to the output of Gepasi, the two systems respond the same. Output from Gepasi and Matlab is listed in chapter 4.

### *3.8 Summary*

This chapter looked at previously developed methods for determining the steady state behavior for systems of autonomous ordinary differential equations. These methods included defining stability requirements for linear models as well as linearizing nonlinear models to test for stability in local regions. It then used these methods to analyze chemical systems, such as the Brusselator and Schnackenberg's model, by reformulating the chemical models into systems of differential equations. A stability analysis was performed on these stable states to determine the sensitivity of the steady state behavior under different parameter settings. By considering interesting chemical systems with bifurcations, a variety of different types of behavior were described. Finally, by evaluating models of transcription and translation, it has been shown that these interesting reaction models can be incorporated into biological systems to define specific aspects of biological behavior. With a better understanding of different forms of stability, such as asymptotic and oscillatory stability, and different techniques for finding that stability, to include linearization techniques and substitutions, it will be easier to construct models that when incorporated together, represent real biological systems that can be tested to fulfill Air Force needs.

# *IV. Model Results*

## *4.1 Overview*

A number of the models constructed in Chapter 3, the analysis portion of this thesis, potentially exhibit interesting behavior such as limit cycle solutions and multiple stable steady state solutions. Under specific conditions, this behavior may be sensitive to variable concentrations or model parameters. This Chapter will investigate changes of the small linear system model of Chapter 3, the Brusselator model [27], the Schnackenberg model [30], the model constructed by Hasty et al. [14], a reaction inside the Glucose metabolic pathway, and the larger Glucose model [34] as a result of perturbations of initial conditions and parameters. In this chapter, all quantities are nondimensional.

## *4.2 A small linear system*

The small system constructed in Chapter 3, namely

$$\begin{aligned} \frac{dX}{dt} &= -(3+\varepsilon)X + 3Y \\ \frac{dY}{dt} &= \frac{(2+\varepsilon)^2}{4}X - Y, \end{aligned} \tag{4.1}$$

was analyzed to find the values of the perturbation term $\varepsilon$ where steady state behavioral changes take place. In this Chapter, a numerical analysis of specific cases adds more detail to how the concentrations of $X$ and $Y$ behave given different values for $\varepsilon$. The basic analysis of this linear system leads to a greater understanding of how a linearized system will behave in a neighborhood of a steady state.

The analysis of Chapter 3 shows that any value $\varepsilon$ belonging to the interval $(-8/3, 0)$ will lead to a stable solution. Figure 4.1 represents $X$ and $Y$ when $\varepsilon$ is given the value $\varepsilon = -1$, and the initial conditions for $X$ and $Y$ are $X_0 = Y_0 = 5$. The eigenvalues of the system under these conditions are $\lambda_1 = -.5$, and $\lambda_2 = -2.5$

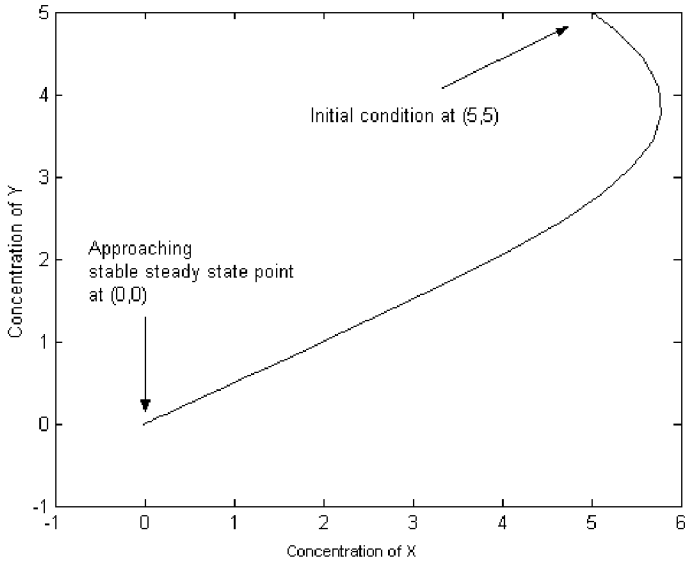


Figure 4.1. Steady state stability in a linear system when $\varepsilon = -1$

which implies the solution is stable as predicted in Chapter 3. With 50 time steps, the final concentration values for $X$ and $Y$ reach the steady state value of $X = 0$ and $Y = 0$ asymptotically.

The analysis of Chapter 3 also shows that any positive value for $\varepsilon$ will lead to an unstable solution. Figure 4.2 represents $X$ and $Y$ when $\varepsilon$ is given the value $\varepsilon = 1$, and the initial conditions for $X$ and $Y$ are $X_0 = Y_0 = 5$. The eigenvalues of the system are $\lambda_1 = .5$, and $\lambda_2 = -5.5$ which implies the solution is a saddle. With 5 time steps, the final concentration values for $X$ and $Y$ moves quickly away from the steady state value to a concentration of $X = 45.8812$ and $Y = 68.8218$.

This saddle point behavior is seen again in when $\varepsilon < -\frac{8}{3}$. To illustrate this, $\varepsilon$ was given the value $\varepsilon = -2.9$, with the initial conditions for $X$ and $Y$ left at $X_0 = Y_0 = 5$. The eigenvalues under these conditions are $\lambda_1 = -1.45$, and $\lambda_2 = 0.35$ which again implies the solution is a saddle as predicted in Chapter 3. With 5 time steps, the final concentration values for $X$ and $Y$ moves away from the steady state

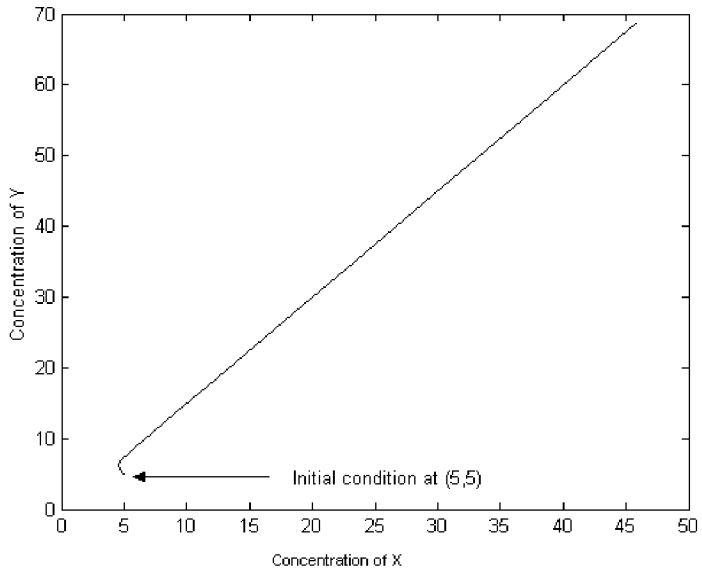


Figure 4.2. Steady state instability in a linear system when $\varepsilon = 1$

value to a concentration of $X = 69.6323$ and $Y = 10.4475$. This is plotted in Figure 4.3.

These models show both a case of asymptotic stability to a steady state, and two cases of divergent behavior in the form of a saddle point. They also indicate that solutions of a system can be highly sensitive to small perturbations of parameters. This analysis gives a partial description of the behavioral characteristics of a linear system. This analysis can then be expanded to evaluate more complex systems through the process of linearization.

### *4.3 The Brusselator*

In equations (4.2), the differential equations representing the Brusselator equation include parameters that can be perturbed just as the previous linear model. These parameters in the Brusselator, represented as $a$ and $b$, are controlled values that, when perturbed, can introduce very interesting behavior to the system. This

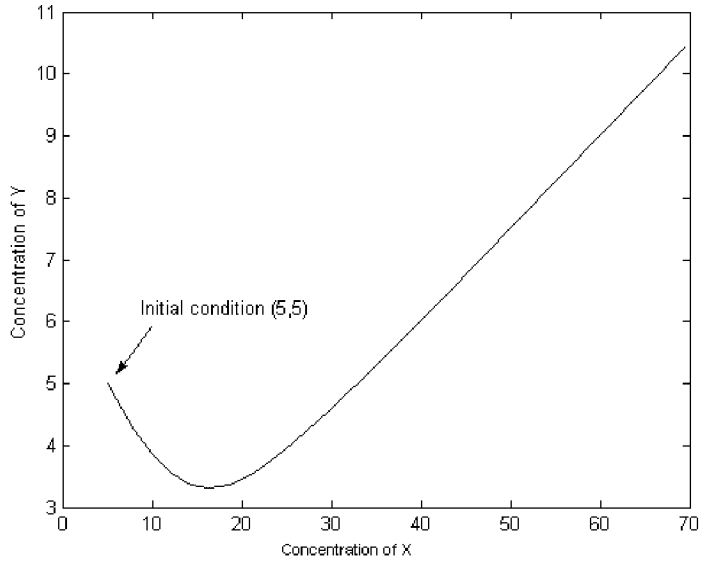


Figure 4.3. Steady state instability in a linear system when $\varepsilon = -2.9$

section will analyze the behavior of the Brusselator to distinguish the difference in behavior of a nonlinear systems from that of linear systems.

$$\begin{aligned} \frac{dx}{dt} &= a - (b+1)x + x^2 y \\ \frac{dy}{dt} &= bx - x^2 y \end{aligned} \tag{4.2}$$

In Figure 4.4, $a = 0.60710$, and $b = 1.5$, so the eigenvalues, based on the linearization of section 3.3.2, are $\lambda_{1,2} = 6.571\,1 \times 10^{-2} \pm .603\,54i$, which implies cyclic divergence from the steady state point located at $(x, y) = (0.6071, 2.470\,7)$. However, once outside a small neighborhood of the steady state, the concentrations approach a limit cycle. This is not unexpected since linearization indicates only local behavior.

As another example, consider the case, $a = 0.8071$ and $b = 1.5$, so the eigenvalues are $\lambda_{1,2} = -7.570\,5 \times 10^{-2} \pm .803\,54i$, which implies the steady state located

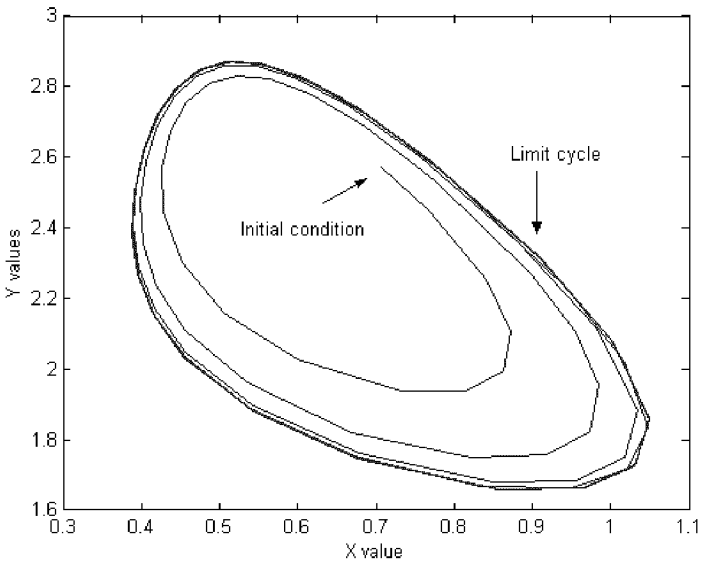


Figure 4.4. Limit cycle behavior in the Brusselator

at $(x, y) = (0.8071, 1.8585)$ is an oscillatory stability point. In Figure 4.5, the initial concentrations are located at $(x_0, y_0) = (0.9071, 1.9585)$, and are therefore within a sufficiently small neighborhood of the steady state point to allow linearization to be used as a means of determining the concentrations behavior. As the concentrations change with time, they approach the steady state.

Although the Brusselator model is a hypothetical chemical oscillator, the behavior within it may exist inside larger systems that are models of cellular functions. By understanding how linearity affects this system, more awareness can be gained about the conditions linearity imposes on larger systems. This may eventually lead to a better understanding of large models involving cellular processes.

## *4.4 Schnackenberg*

The Schnackenberg model, presented in Chapters 2 and 3, has stability characteristics that are important to understand when developing large biological models. This section will do a sensitivity analysis on both a limit cycle and stable steady

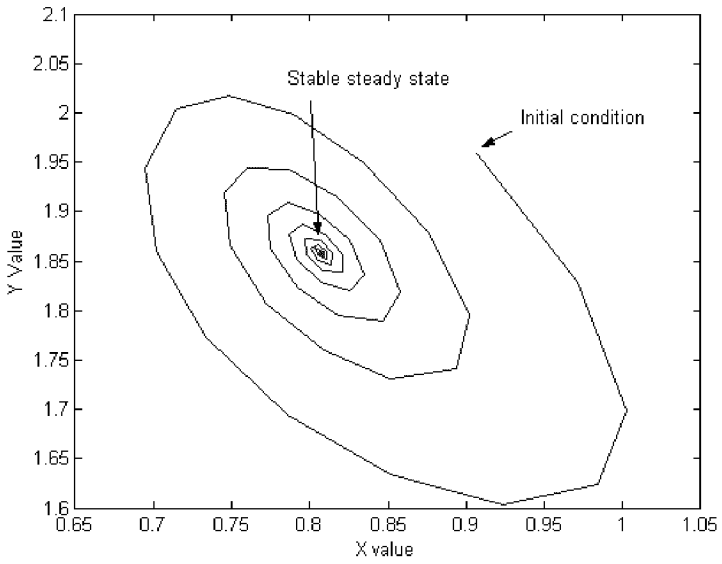


Figure 4.5. Cyclic stability in the Brusselator

state within the Schnackenberg model. The sensitivity analysis will consist of perturbations of reactant concentrations from stable steady state and limit cycle as well as perturbations of constant parameters within the model. Each sensitivity analysis is evaluated by a plot showing how concentrations change through time. In these plots, the concentration of X is measured on the x-axis, and the concentration of Y is measured on the y-axis. This analysis will lead to understand how systems may be very sensitive to small perturbations.

Figure 4.6 indicates limit cycle behavior for the Schnackenberg model given by equation (3.49) and (3.50). This model's constant parameter values are $a = 0.2091$ and $b = .1$. From equations (3.53) and (3.54) the model has a steady state located at the point $(X, Y) = (0.3091, 2.1884)$. From the characteristic equation (3.56), the eigenvalues have a positive real part, so the steady state is unstable. The initial condition for Figure 4.6 is $(X_0, Y_0) = (0.4091, 2.2884)$. This leaves the initial condition within the interior of the limit cycle which is indicated in Figure 4.6. If the initial point is any point inside the limit cycle, other than the steady state point, the resulting trajectory will trace a path out to the limit cycle.

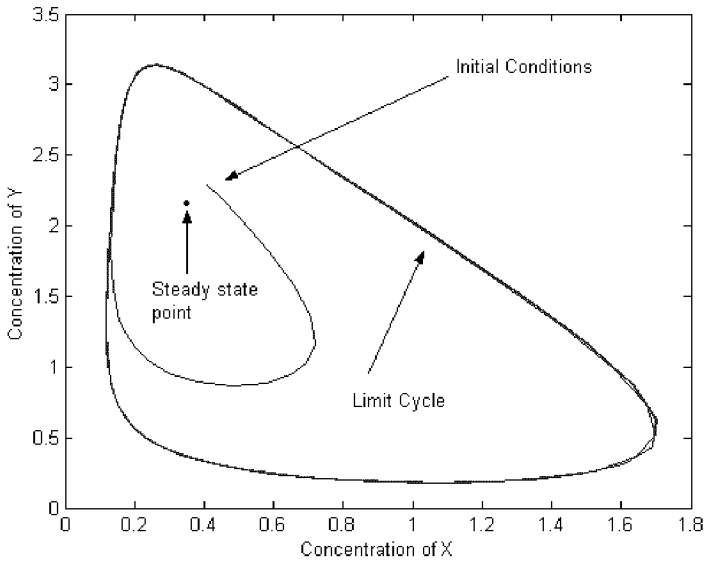


Figure 4.6. Interior convergence to a limit cycle in a Schnackenberg model

Figure 4.7 shows the same limit cycle with initial condition $(X_0, Y_0) = (1.3091, 3.1884)$. This puts the initial condition outside the limit cycle. Note that the trajectory which the solution traces moves away from the limit cycle before it moves to it.

Figure 4.8 shows another type of behavior for the Schnackenberg model. This model's constant parameter values $a = 0.1091$ and $b = 0.1$. From equations (3.53) and (3.54) the model has a steady state located at the point $(X, Y) = (0.2091, 2.4952)$. From equation (3.56), the eigenvalues have negative real parts so the steady state is stable. The initial condition in this model is $(X_0, Y_0) = (0.3091, 2.5952)$. This leaves the initial condition only a small distance from the stable steady state point. The concentration moves much further from the steady state before returning back. As the concentration moves back towards steady state, its trajectory wraps tightly around the steady state point. To graphically observe convergence to the steady state point, every tenth and then every 50th point was plotted.

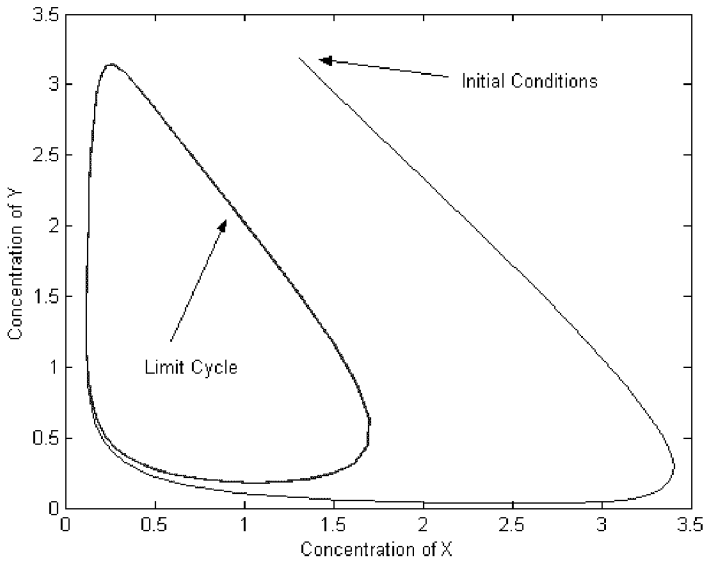


Figure 4.7. Exterior convergence to a limit cycle in a Schnackenberg model

In Figure 4.8, the graph shows that the system is very sensitive to small perturbations. In a biological model, this type of behavior might be detrimental to a cell. In Figure 4.8, the small perturbation leads the concentrations of the reactants to toxic levels that may kill or damage the organism. In this case, it becomes very important for the larger biological model to regulate this process so as to prevent its functions from loosing control. On the other hand, this type of behavior may be integrated into a cell model as an amplifier that triggers a switch in later reactants from one steady state concentration to another. This type of behavior is described in the next section.

The final Schnackenberg example shows stable behavior (Figure 4.9). This model's constant parameter values are $a = 0.0091$ and $b = 0.1$. The model has a stable steady state located at the point $(X, Y) = (0.1091, 0.7679)$. The initial condition in this model is $(X_0, Y_0) = (0.2091, 0.8679)$.

The stability in this example is asymptotic which is quite different from the oscillatory convergence or the convergence to a limit cycle of the previous examples.

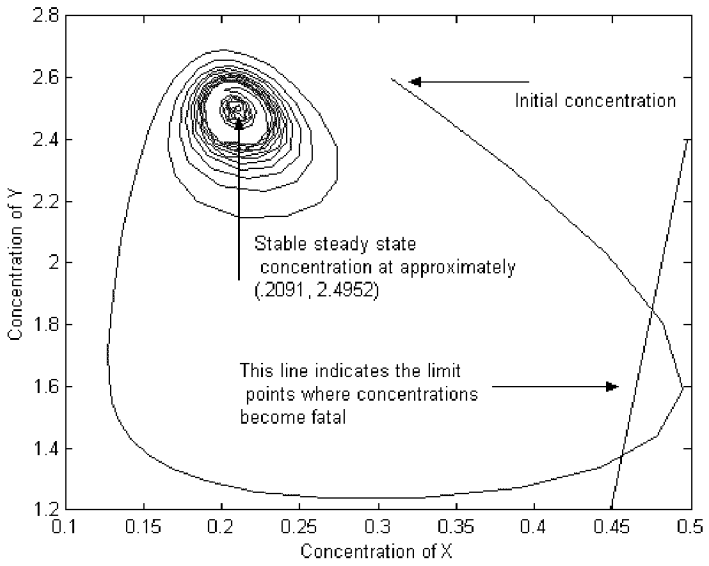


Figure 4.8. Oscillator convergence in the Schnackenberg model

In this and the previous examples, the steady state point is based on the constant parameter values using equations (3.53) and (3.54). Furthermore, each example had an initial condition placed at $(X_0, Y_0) = (X_{ss} + \varepsilon_1, Y_{ss} + \varepsilon_2)$ where $\varepsilon_1 = \varepsilon_2 = 0.1$. The only perturbation in each model came from the model's constant parameter $a$. Therefore, the behavior of the Schnackenberg model is obviously sensitive to perturbations from the constant parameter values.

The large behavioral changes due to small perturbations of parameters in the small Schnackenberg model give insight into the great complexity of larger systems. From the Schnackenberg model, it becomes very clear that larger biological systems need to control the range of perturbation on their parameter values so as to prevent extreme changes in the system that could kill the organism.

### *4.5 Hasty et al.'s Small Model*

Hasty et al.'s model, developed in Chapters two and three, was an example of a bistable system. In this section, Figures 4.10 and 4.11, representing a stability

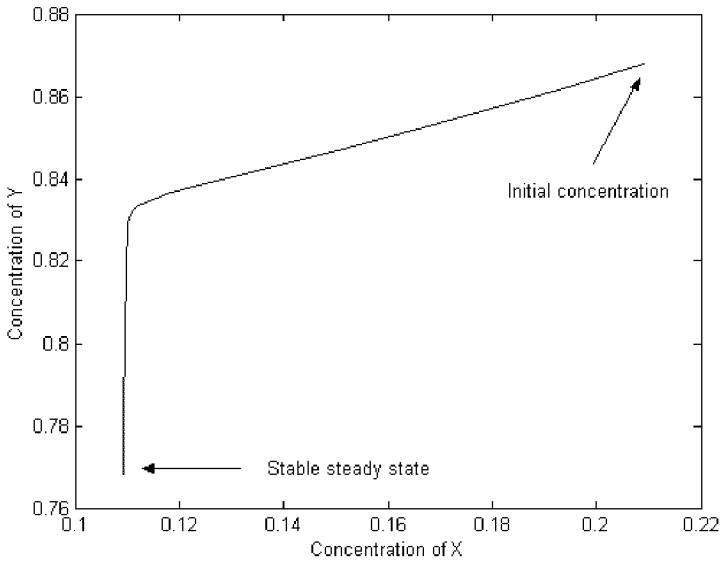


Figure 4.9. Asmptotic convergence in the Schnackenberg model

analysis, will give insight into exactly how the system behaves under small perturbations in the concentrations of reactants $X$ and $X_2$ from their steady state. Figures 4.12 and 4.13, representing bifurcation diagrams, will show exactly how the steady state concentration of the reactant $X$ behaves under different values of the constant parameters $\alpha$ and $\delta$. Using both sensitivity graphs and bifurcation diagrams, the behavior of Hasty et al.'s model will be fully understood.

Figure 4.10 shows stable behavior. Its constant parameter values are $\alpha = 50$, $\delta = 0.4$, $K_1 = 1$, $K_2 = 1$, $\beta = K_1K_2$, $\gamma = 10$, and $dT = 10$. For these parameter values, the model has a three steady states located at the points $(X, X_2)_1 = (0.0552, 0.0030)$, $(X, X_2)_2 = (0.1499, 0.0225)$, $(X, X_2)_3 = (4.8349, 23.3761)$ where $X$ is respectively one of the roots of equation (3.99) and $X_2$ satisfies equation (3.91) with $\frac{k_1}{k_{-1}} = K_1 = 1$. The initial condition in this model is $(X, X_2)_0 = (0.0652, 0.0130)$. Consequently, the initial condition is a small distance from the steady state point $(X, X_2)_1$. As is observed in Figure 4.10, the concentrations quickly return to the original steady state. Thus, it is clear that the steady state at $(X, X_2)_1$ is stable.

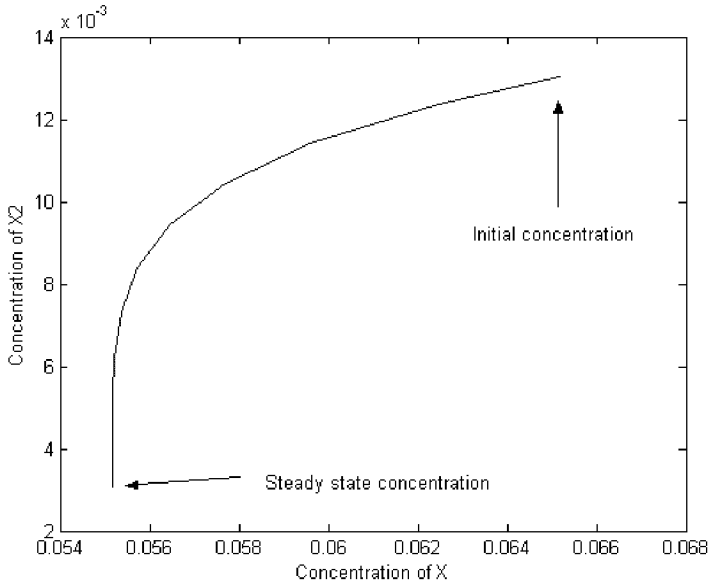


Figure 4.10. Lower steady state convergence in the Hasty model

Figure 4.11 also shows stable behavior. All parameter values are the same as for the previous example. The only change is that $(X, X_2)_0 = (0.1552, 0.1030)$. In fact, this puts steady state $(X, X_2)_2$ between $(X, X_2)_1$ and initial condition. As is observed in Figure 4.11, the concentration moves away from both steady state $(X, X_2)_1$ and steady state $(X, X_2)_2$, and converges to steady state $(X, X_2)_3$. This indicates that the steady state at $(X, X_2)_2$ is unstable, and steady state $(X, X_2)_3$ is stable. Thus, these numerical examples are consistent with the analysis by Hasty et al. concluding that this is a bistable system.

The next two figures are bifurcation diagrams that explain the behavior of the steady state concentrations as a function of a single parameter. Figure 4.12 describes the steady state concentration of X as a function of $\alpha$. The constant parameter values are $\delta = 0.4$, $K_1 = 1$, $K_2 = 1$, $\beta = K_1 K_2$, $\gamma = 10$, and $dT = 10$. For each $\alpha$, the values of $[X]$ are determined by solving equation (3.99) where $u = [X]$. The constant parameter $\alpha$ may change for a number of reasons. Recall from equations (3.95) and (3.96) that $\alpha = nk_3 p_0 \beta dT$, so if the concentration of RNA polymerase $p_0$ were to somehow change, $\alpha$ would change in proportion to it. As indicated previously, when

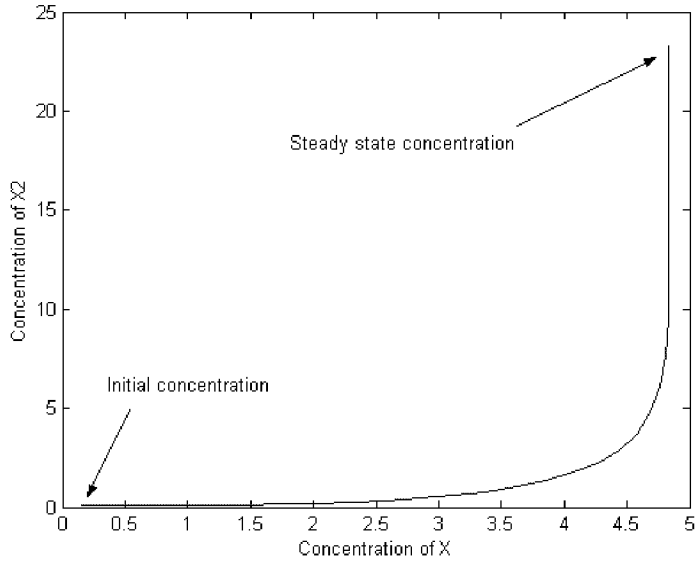


Figure 4.11. Upper steady state convergence in the Hasty model

$\alpha = 50$, the steady state points are at approximately $(X, X_2)_1 = (0.0552, 0.0030)$, $(X, X_2)_2 = (0.1499, 0.0225)$, and $(X, X_2)_3 = (4.8349, 23.3761)$. Also, by the stability analysis, it has been show which steady states are stable and which are unstable. The graph shows that bifurcation points exist at approximately $\alpha = 20$, and $\alpha = 63$. At $\alpha = 20$, two of the steady state solutions go from complex to real values and then at $\alpha = 63$ two steady states go from real to complex values. Thus when $\alpha \lesssim 20$, there is one real steady state value, when $20 \lesssim \alpha \lesssim 63$, there are three real steady state values, and again when $63 \lesssim \alpha$, there is one real steady state value. Figure 4.12 is exactly the same as Figure 1a of Hasty et al.

This bifurcation plot can also be associated with $\delta$. Recall from equation (3.74) that $\delta$ is the basal rate of production of $X$. This may change depending on occurrences outside of the system. For $\delta$ varying about the x-axis, $\alpha = 25$, and all other parameters held constant with rates described previously, it is obvious that small changes in $\delta$ lead to large changes in the system. For example, at $\delta = .4$, there are three stable states. Yet, by changing $\delta$ such that $\delta = 1.4$, there is only one real steady state. This implies that $\delta$ is a very sensitive parameter.

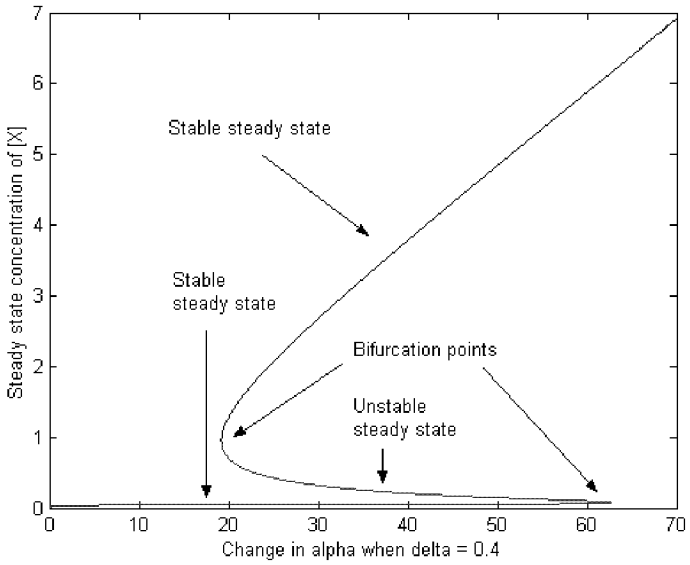


Figure 4.12. A Hasty model bifurcation plot that maps the steady state of [X] with respect to $\alpha$

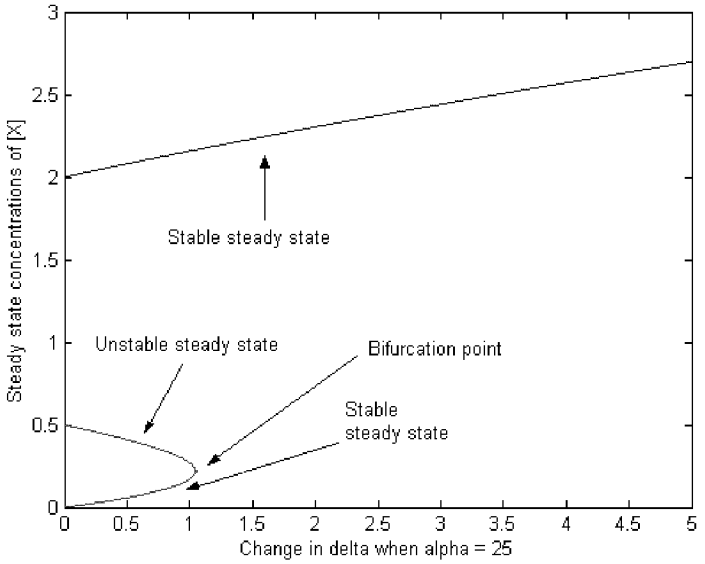


Figure 4.13. A Hasty model bifurcation plot that maps the steady state of [X] with respect to $\delta$

Bistability in Hasty et al.'s model shows how perturbations can permanently affect a system. Furthermore, through stability analysis, the behavior of each steady state can be described, and through bifurcation diagrams, these steady states can be tracked with different concentrations of parameters. In larger models, bistability may act as a regulatory switch, to control the model's behavior.

## *4.6 Glucose*

*4.6.1 Small Glucose Model.* An evaluation on a small subsystem within the Glucose model showed interesting behavior. The subsystem evaluated was the reaction taking $\alpha$-D-Glucose-6-phosphate to $\beta$-D-Fructose-6-phospate in the presence of Phosphoglucoisomerase or the reaction taking 3-Phosphoglycerate to 2-Phosphoglycerate in the presence of Phosphoglycerate Mutase. The reaction equations for the first small reaction system are

$$\alpha\text{-D-Glucose-6-phosphate } (y_1) + \text{Phosphoglucoisomerase } (y_2) \underset{k_5}{\overset{k_4}{\rightleftharpoons}} \text{Catalyst Transition 2 } (y_3)$$

$$\text{Catalyst Transition 2 } (y_3) \underset{k_7}{\overset{k_6}{\rightleftharpoons}} \beta\text{-D-Fructose-6-phospate } (y_4) + \text{Phosphoglucoisomerase } (y_2).$$

This leads to differential equations of the form

$$\begin{aligned} \frac{dy_1}{dt} &= -k_4 y_1 y_2 + k_5 y_3 \\ \frac{dy_2}{dt} &= -k_4 y_1 y_2 + (k_5 + k_6) y_3 - k_7 y_2 y_4 \\ \frac{dy_3}{dt} &= k_4 y_1 y_2 - (k_5 + k_6) y_3 + k_7 y_2 y_4 \\ \frac{dy_4}{dt} &= k_6 y_3 - k_7 y_2 y_4 \end{aligned} \tag{4.3}$$

since, $\frac{dy_2}{dt} + \frac{dy_3}{dt} = \frac{d(y_2+y_3)}{dt} = 0$, $y_2 + y_3 = c$ where $c = y_2(0) + y_3(0)$. Thus, $y_3 = c - y_2$. Substituting into (4.3) results in the equations

$$\frac{dy_1}{dt} = -k_4 y_1 y_2 + k_5(c - y_2) \tag{4.4}$$

$$\frac{dy_2}{dt} = -k_4 y_1 y_2 + (k_5 + k_6)(c - y_2) - k_7 y_2 y_4 \tag{4.5}$$

$$\frac{dy_4}{dt} = k_6(c - y_2) - k_7 y_2 y_4. \tag{4.6}$$

Steady state values can be located by setting $\frac{dy_1}{dt} = 0$, $\frac{dy_2}{dt} = 0$, and $\frac{dy_4}{dt} = 0$. The resulting equations are

$$0 = -k_4 y_1 y_2 + k_5(c - y_2) \tag{4.7}$$

$$0 = -k_4 y_1 y_2 + (k_5 + k_6)(c - y_2) - k_7 y_2 y_4 \tag{4.8}$$

$$0 = k_6(c - y_2) - k_7 y_2 y_4 \tag{4.9}$$

Let $y_2 = \alpha$ and substitute into equations (4.7)-(4.9).

$$0 = -k_4 y_1 \alpha + k_5(c - \alpha) \tag{4.10}$$

$$0 = -k_4 y_1 \alpha + (k_5 + k_6)(c - \alpha) - k_7 \alpha y_4 \tag{4.11}$$

$$0 = k_6(c - \alpha) - k_7 \alpha y_4 \tag{4.12}$$

Solve equations (4.10) and (4.12) for $y_1$and $y_4$ to get steady state concentrations in terms of $\alpha$.

$$y_1 = \frac{k_5(c-\alpha)}{k_4\alpha} \tag{4.13}$$

$$y_4 = \frac{k_6(c-\alpha)}{k_7\alpha} \tag{4.14}$$

Equation (4.11) is a linear combination of equations (4.10) and (4.12) so $\alpha$ remains an arbitrary value. Therefore, the steady state values for $y_1$,$y_2$ ,$y_3$ , and $y_4$ are given by

$$y_1 = \frac{k_5(c-\alpha)}{k_4\alpha} \tag{4.15}$$

$$y_2 = \alpha \tag{4.16}$$

$$y_3 = c-\alpha \tag{4.17}$$

$$y_4 = \frac{k_6(c-\alpha)}{k_7\alpha}. \tag{4.18}$$

Figure 4.14 describes the steady state concentration of $\alpha$-D-Glucose-6-phosphate ($y_1$), and $\beta$-D-Fructose-6-phospate ($y_4$) as a function of the steady state concentration of Phosphoglucoisomerase ($y_2$). In this graph, $c = 100$, $k_4 = 2$, $k_5 = 1$, $k_6 = 2$, $k_7 = 1$, and $\alpha$ lies within the range [25, 75]. As values for $\alpha$ change, a steady state line is created that defines the steady state concentrations for the system. As observed from the equations (4.15)-(4.18), different values for the kinetic constants lead to different shapes for the steady state curve. In the next section, the small model will be incorporated into a large model of the glucose pathway. An example which is consistent with incorporating the small model into the large model includes initial conditions $y_1 = 50$, $y_2 = 100$, $y_3 = 0$, and $y_4 = 0$. This yields a steady state value of $y_1 = 0.456$, $y_2 = 52.281$, $y_3 = 47.718$, and $y_4 = 1.825$.

Testing the steady state behavior of the steady state curve under different values of the kinetic constants showed that the steady state curve was stable. That

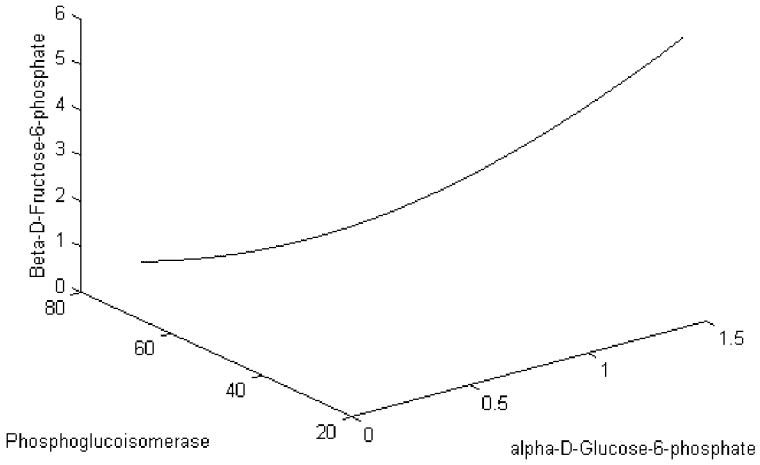


Figure 4.14. Steady state curve for the small glucose model

is, values that were pushed from the steady state line, returned to it. Unfortunately, tracking the system to a steady state point is more difficult than simply substituting in an initial concentration $\alpha$ for Phosphoglucoisomerase ($y_2$). If a concentration within the system is moved away from its steady state point on the steady state curve, then, in the process of moving back to the steady state curve, $\alpha$ changes. Thus, the system will move to a new point on the steady state curve. So, the system is stable with respect to the steady state curve, but not with respect to individual steady state points. This is indicated in Figure 4.15.

Further analysis on the small system determined whether there was different types of behavior near the steady state for different combinations of positive real kinetic constants.

First, the system (4.3) was simplified to (4.4)-(4.6) where the third equation of system (4.3) was omitted because $y_3 = c - y_2$. The Jacobian of the system (4.4)-(4.6) takes the form

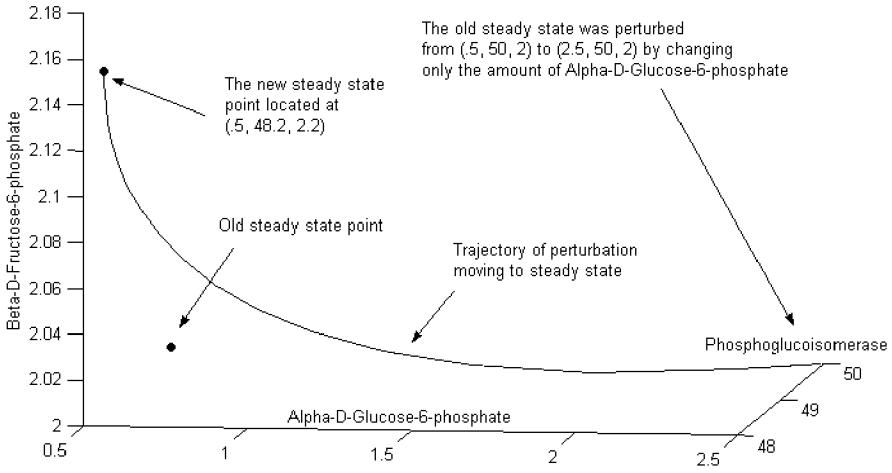


Figure 4.15. Convergence to a new point on the steady state curve in the small glucose model

$$J = \begin{pmatrix} -k_4 y_2 & -k_4 y_1 - k_5 & 0 \\ -k_4 y_2 & -k_4 y_1 - (k_5 + k_6) - k_7 y_4 & -k_7 y_4 \\ 0 & -k_6 - k_7 y_4 & -k_7 y_4 \end{pmatrix}. \tag{4.19}$$

Then solving for the characteristic equation, $\det|\lambda I - J| = 0$

$$\det \begin{pmatrix} \lambda + k_4 y_2 & k_4 y_1 + k_5 & 0 \\ k_4 y_2 & \lambda + k_4 y_1 + (k_5 + k_6) + k_7 y_4 & k_7 y_4 \\ 0 & k_6 + k_7 y_4 & \lambda + k_7 y_4 \end{pmatrix} = 0 \tag{4.20}$$

yields the cubic equation

$$\lambda^3 + (2k_7 y_4 + k_4 y_1 + k_5 + k_6 + k_4 y_2)\lambda^2 + a_0 \lambda = 0.$$

where $a_0 = k_4 y_1 k_7 y_4 + k_5 k_7 y_4 + 2k_4 y_2 k_7 y_4 + k_4 y_2 k_6$ which simplifies to a quadratic by factoring out a $\lambda$. Thus the first root is

$$\lambda = 0 \tag{4.21}$$

and the quadratic equation is

$$\lambda^2 + (2k_7y_4 + k_4y_1 + k_5 + k_6 + k_4y_2)\lambda + a_0 = 0 \tag{4.22}$$

where $a_0 = k_4y_1k_7y_4 + k_5k_7y_4 + 2k_4y_2k_7y_4 + k_4y_2k_6$.

The solution to the quadratic takes the form $\lambda = -k_7y_4 - \frac{1}{2}k_4y_1 - \frac{1}{2}k_5 - \frac{1}{2}k_6 - \frac{1}{2}k_4y_2 \pm \frac{1}{2}\sqrt{D_I}$, where $D_I = 2k_4^2y_1y_2 + 4k_7y_4k_6 - 4k_4y_2k_7y_4 + 4k_7^2y_4^2 - 2k_4y_2k_6 + 2k_4y_1k_6 + 2k_4y_1k_5 + 2k_5k_4y_2 + k_4^2y_1^2 + k_5^2 + 2k_5k_6 + k_6^2 + k_4^2y_2^2$

Substituting the steady state points in for $y_1$, $y_2$, and $y_4$ given by equations (4.15), (4.16), and (4.18) and simplifying gives the value $\lambda = -\frac{1}{2\alpha}(2k_6c - k_6\alpha + k_5c + k_4\alpha^2) \pm \frac{1}{2\alpha}\sqrt{r_1}$ where $r_1 = 2k_4k_5\alpha^2c + k_6^2\alpha^2 + 2k_4\alpha^3k_6 + 4k_6^2c^2 + 2k_5k_6\alpha c + k_5^2c^2 + k_4^2\alpha^4 - 4k_6^2\alpha c - 4k_4k_6\alpha^2c$

If $r_1 < 0$, then the system has oscillatory behavior. Now, $r_1$ can be simplified by omitting some positive terms. This can be accomplished by setting $k_5 = 0$. The new equation is

$$k_6^2\alpha^2 + 2k_4k_6\alpha^3 + 4k_6^2c^2 + k_4^2\alpha^4 - 4k_6^2\alpha c - 4k_4k_6\alpha^2c < 0 \tag{4.23}$$

This equation then factors to the form

$$\left(k_6\alpha + k_4\alpha^2 - 2k_6c\right)^2 < 0 \tag{4.24}$$

which is a contradiction to

$$\left(k_6\alpha + k_4\alpha^2 - 2k_6c\right)^2 \geq 0. \tag{4.25}$$

Thus, the system has no oscillatory behavior.

The value for $\lambda$ comes from equation (4.22), by way of the quadratic formula $\lambda_{\pm} = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$. If $4ac < 0$, for values of $\alpha$ or the kinetic constants, then $\lambda_+ > 0$ and the steady state curve has divergent behavior. Now $4ac = 4(k_4k_7y_1y_4 + k_5k_7y_4 + 2k_4k_7y_2y_4 + k_4k_6y_2)$ from equation (4.22). Substituting the steady state points in for $y_1$, $y_2$, and $y_4$ given as equations (4.15), (4.16), and (4.18) and simplifying gives the value $4ac = 4k_6\frac{k_5c^2 - k_5c\alpha + 2k_4\alpha^2c - k_4\alpha^3}{\alpha^2}$. Now if $k_5c^2 - k_5c\alpha + 2k_4\alpha^2c - k_4\alpha^3 < 0$, then there exists divergent behavior. By factoring, $k_5c(c-\alpha) + k_4\alpha^2(2c-\alpha) < 0$, but from equation (4.17), $c > \alpha$,for all values of $c$ and $\alpha$, so the steady state curve is stable for all $c$ and $\alpha$.

The zero eigenvalue, presented in equation (4.21), introduces the possibility that the system represented by equations (4.4)-(4.6) may be reduced to a system of two equations. Reducing the system can provide further understanding of the models steady state behavior.

Note from equations (4.4)-(4.6), that $\frac{dy_2}{dt} - \frac{dy_1}{dt} = \frac{dy_4}{dt}$. This implies that $\frac{dy_1}{dt} + \frac{dy_4}{dt} - \frac{dy_2}{dt} = 0$, or more appropriately that $y_4 - y_2 + y_1 = c_2$ where

$$c_2 = y_4(0) - y_2(0) + y_1(0). \tag{4.26}$$

By substituting $y_4 = c_2 + y_2 - y_1$, equations (4.4) and (4.6) become

$$\frac{dy_1}{dt} = -k_4y_1y_2 + k_5(c - y_2) \tag{4.27}$$

$$\frac{dy_2}{dt} = -k_4y_1y_2 + (k_5 + k_6)(c - y_2) - k_7y_2(c_2 + y_2 - y_1) \tag{4.28}$$

At steady state, values can be located by setting $\frac{dy_1}{dt} = 0$ and $\frac{dy_2}{dt} = 0$. The resulting equations are

$$0 = -k_4 y_1 y_2 + k_5(c - y_2) \tag{4.29}$$

$$0 = -k_4 y_1 y_2 + (k_5 + k_6)(c - y_2) - k_7 y_2(c_2 + y_2 - y_1) \tag{4.30}$$

Solve equations (4.29) and (4.30) for $y_1$and $y_4$ to get the steady state concentrations

$$y_1 = -\frac{k_5}{k_4}\left(\frac{2ck_7k_4 + k_4k_6 + k_7c_2k_4 + k_7k_5 - \sqrt{r_2}}{k_4k_6 + k_7c_2k_4 + k_7k_5 - \sqrt{r_2}}\right) \tag{4.31}$$

$$y_2 = \frac{1}{2k_7k_4}\left(-k_4k_6 - k_7c_2k_4 - k_7k_5 + \sqrt{r_2}\right) \tag{4.32}$$

where $r_2 = k_4^2k_6^2 + 2k_4^2k_6k_7c_2 + 2k_4k_6k_7k_5 + k_7^2c_2^2k_4^2 + 2k_7^2c_2k_4k_5 + k_7^2k_5^2 + 4k_7k_4^2k_6c + 4k_7^2k_4k_5c$. Equations (4.31) and (4.32) involve no free parameters, so by reducing the system of equations (4.4)-(4.6) to two equations given by (4.27) and (4.28), the steady state line becomes a steady state point. Plugging in the initial conditions $y_1(0) = 50$, $y_2(0) = 100$, $y_3(0) = 0$ , $y_4(0) = 0$ so that $c = y_2(0) + y_3(0) = 100$ and $c_2 = \ y_4(0) - y_2(0) + y_1(0) = -50$ where $k_4 = 2$, $k_5 = 1$, $k_6 = 2$, and $k_7 = 1$, gives $r_2 = 13\,025$, such that $y_1 = 0.456\,36$, $y_2 = 52.282$, $y_3 = 47.718$, and $y_4 = 1.825\,6$. These results are consistent with the results produced previously by the three equation system.

The Jacobian of the system (4.27)-(4.28) takes the form

$$J = \begin{pmatrix} -k_4y_2 & -k_4y_1 - k_5 \\ -k_4y_2 + k_7y_2 & -k_4y_1 - k_5 - k_6 - k_7c_2 - 2k_7y_2 + k_7y_1 \end{pmatrix}. \tag{4.33}$$

Then solving for the characteristic equation, $\det|\lambda I - J| = 0$

$$\det\begin{pmatrix} \lambda + k_4y_2 & k_4y_1 + k_5 \\ k_4y_2 - k_7y_2 & \lambda + k_4y_1 + k_5 + k_6 + k_7c_2 + 2k_7y_2 - k_7y_1 \end{pmatrix} = 0 \tag{4.34}$$

gives the polynomial $\lambda^2 + b_1\lambda + b_0 = 0$. where $b_1 = k_4y_1 + k_5 + k_6 + k_7c_2 + 2k_7y_2 - k_7y_1 + k_4y_2$, and $b_0 = k_4k_6y_2 + k_7c_2k_4y_2 + 2k_7k_4y_2^2 + k_7k_5y_2$. The eigenvalues are given by $\lambda = \frac{1}{2}(-b_1 \pm \sqrt{(b_1)^2 - 4b_0})$. Now, $\lambda$ is very complicated and although an analytic evaluation may be possible, simplifying $\lambda$ is difficult. One reason for this is due to the negative value within equation (4.26). The negative values allow the possibility for $c_2 < 0$, so $c_2$ cannot be used to simplify $\lambda$. For this reason, a numerical evaluation is more practical. The eigenvalues corresponding to the steady state point evaluated above are $\lambda_+ = -55.967$ and $\lambda_- = -106.61$. The negative eigenvalues correspond to stable behavior within a neighborhood of the steady state point $(y_1, y_2, y_3, y_4)_{ss} = (0.456\,36, 52.282, 47.718, 1.825\,6)$.

This system observed steady state behavior in the form of a stable steady state curve when evaluated as a system of three differential equations and as a stable steady state point when the system was reduced to two equations. The three equation system had a single zero eigenvalue corresponding to a free parameter in the steady state. The zero eigenvalue and free parameter were eliminated when the system was reduced and although the original system consisting of four equations was not evaluated, it is likely that it has two zero eigenvalues corresponding to two free parameters. It is worth noting that the behavior of the system was actually easier to analytically evaluate when reduced to three equations, than when reduced to two equations. Although the system was analyzed to look for different forms of behavior near the steady state curve, specifically oscillatory convergence or divergence, the analysis shows that only asymptotically stable behavior exists within this system for positive real values of kinetic constants and the parameter $\alpha$.

This small Glucose model shows that some systems may exhibit stable behavior over a curve, yet perturbing concentrations within the system lead them to flow to a different concentration on that curve. A system may stay within a region of the steady state curve in the following way. With constant small perturbations of one type, the steady state concentrations may flow down that curve. Another

type of perturbation may cause the concentration to flow back up the curve, thus regulating the system. Hypothetically, when modeling a cell, the steady state curve may contain concentrations that are toxic. A large series of the same type of perturbation could lead the steady state concentrations to these toxic levels and kill the cell. Therefore, the concentrations introduced into this smaller system must be regulated through perturbations that are controlled elsewhere in the larger model.

*4.6.2 Large Glucose Model.* The smaller Glucose model is contained in a larger system that is listed in Appendix A and described in Chapter 3, Section 3.7. This larger model incorporates the smaller system as both the second and eighth of ten total reactions. In this model, reactions one, three, and ten are one way reactions that degrade the reactant concentrations of all previous reactions to a steady state value of zero. Thus, the second reaction, denoting the small glucose model, shows no interesting results when analyzed as a part of the larger model. That is, if reaction three is included, then the steady state of reaction two is zero for all of its substrates. However, since reaction eight is the small model and all reactions between reaction three and reaction ten are two way reactions, if reaction ten is omitted from the system, the large system can be run to steady state and the small Glucose model can be analyzed as a part of the larger model. In this case, reaction eight acts as the smaller model and must cooperate with the two way reactions, four through nine.

Analysis of how the smaller system relates to the larger system takes place in steps. By making some assumptions about the larger glucose model, reaction eight acts just the same as the small glucose model already evaluated. The assumptions that give the larger glucose model the characteristics of the smaller glucose model are as follows. First, by setting the concentrations of all the substrates (including catalyst transition reactants) except 3-phosphoglycerate and 2-phosphoglycerate equal to zero, system changes will be initiated only from concentrations of reactants involved in reaction eight. Then, by setting the concentrations of the enzymes,

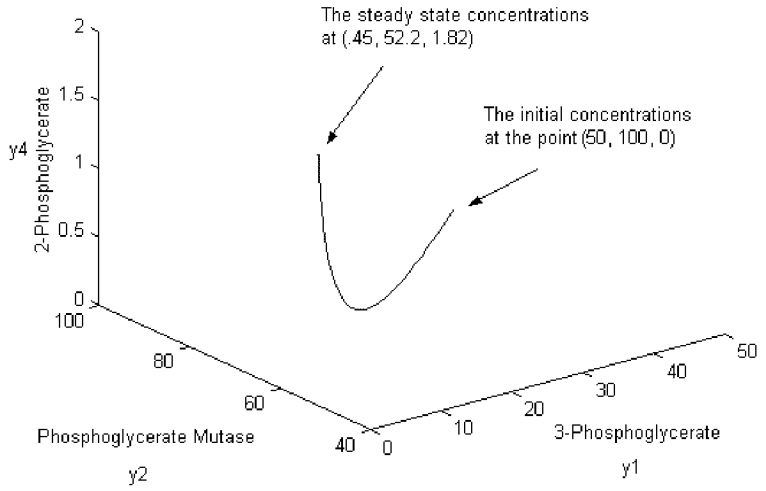


Figure 4.16. The large glucose model with assumptions that make it behave as though it were the small glucose model.

Phosphoglycerate Kinase of reaction seven and Enolase of reaction nine, equal to zero, the reacting species are restricted to only reaction eight. This can also be accomplished by setting the concentration of the coenzyme $Mg^{++}$ equal to zero. The resulting model acts just the same as the small glucose model. Setting the initial concentration of 3-Phosphoglycerate = 50, Phosphoglycerate Mutase = 100, the catalyst transition = 0.00001, and 2-Phosphoglycerate = 0.00001, with forward reaction kinetic constants k27 and k29 equal to 2 and reverse reaction kinetic constants k28 and k30 equal to 1, the systems behavior reaches steady state quickly. The Figure 4.16 indicates how the substrates and enzyme change concentrations through time to reach steady state. Much of the concentration of these reactants is eaten up by the catalyst transition, which at steady state equals 47.59.

By allowing the initial concentration of $\alpha$-D-Glucose to be a nonzero value, and by setting the initial concentrations of the enzymes, Phosphoglycerate Kinase of reaction seven and Enolase of reaction nine, equal to nonzero values, reaction eight now interacts with the larger system. Evaluating this system with small incremental

increases in $\alpha$-D-Glucose, Phosphoglycerate Kinase, and Enolase individually, allows for comparison between the large glucose model and the small glucose model.

To begin, the model was given incremental increases in the concentration of $\alpha$-D-Glucose. With the concentration of Phosphoglycerate Kinase = 1 and the concentration of Enolase = 0. Figure 4.17 describes how concentrations within the small system move to steady state when $\alpha$-D-Glucose = 20 initially. In this case, reaction eight quickly reaches steady state just as it did in the previous system. However, as this is occurring, concentrations of reactants originating from $\alpha$-D-Glucose are flowing through the system. At a later time, these concentrations react to create 3-Phosphoglycerate and 2-Phosphoglycerate. As greater amounts of these substrates are created, a larger amount of the catalyst transition in reaction eight is also created, leading to a depleting of the amount of the enzyme Phosphoglycerate Mutase. A small decrease in 3-Phosphoglycerate also occurs rapidly because the small amount of Phosphoglycerate Kinase introduced into the system, coupled with a small reverse reaction, tries to balance the large concentration of 3-Phosphoglycerate with the small concentration of the substrate, 1,3-Biphosphoglycerate.

The model was then given incremental increases in Phosphoglycerate Kinase, with the concentrations of $\alpha$-D-Glucose = 0 and Enolase = 0. Figure 4.18 shows that as the concentration of Phosphoglycerate Kinase increases, a reverse reaction quickly takes effect in the model by removing concentrations of 3-Phosphoglycerate to balance its concentration with the concentration of 1,3-Bisphosphoglycerate.

Finally, the model was given incremental increases in Enolase, with the concentrations of $\alpha$-D-Glucose = 0 and Phosphoglycerate Kinase = 0. Figure 4.19 shows that as the concentration of Enolase increases, a forward reaction takes effect in the model and removes concentrations of 2-Phosphoglycerate as it is created until a balance in concentration between 2-Phosphoglycerate and Phosphoenolpyruvate

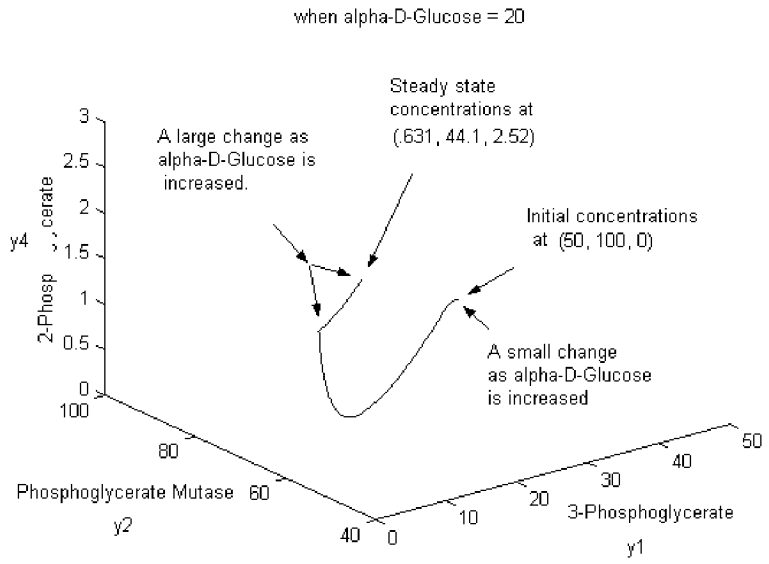


Figure 4.17. Incremental increase in the concentration of $\alpha$-D-Glucose

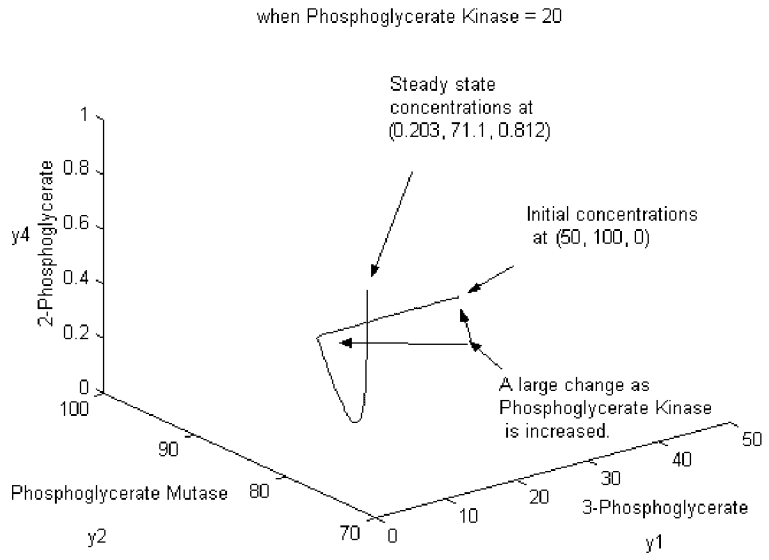


Figure 4.18. Incremental increase in the concentration of Phosphoglycerate Kinase

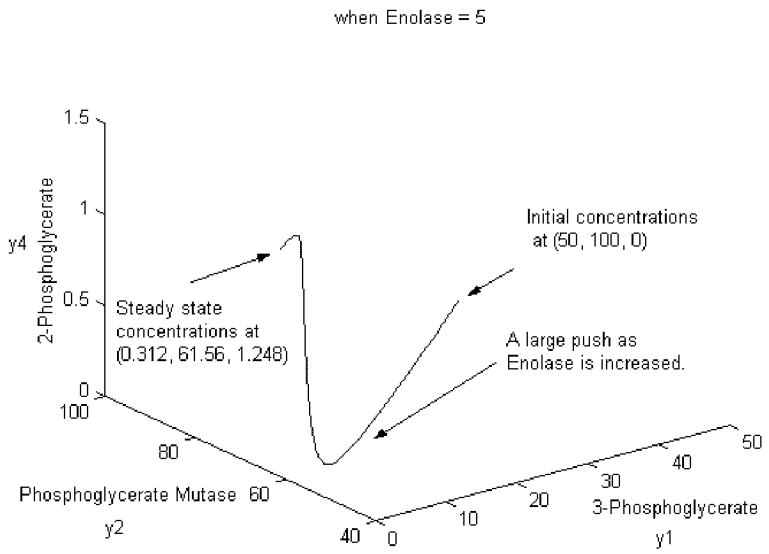


Figure 4.19. Incremental increase in the concentration of Enolase

is created. Thus, more 3-Phosphoglycerate and Phosphoglycerate Mutase must be used up before measurable values of 2-Phosphoglycerate begin to appear.

Analysis of the larger system, omitting reaction ten, shows similarities to the smaller Glucose model. In fact, for the values checked though simulation, the steady state curve for the larger glucose model is virtually equal to the one associated with the small glucose model. In the smaller system, if a small perturbation is created, the model will not return to the original steady state point, but will go to a new steady state point on the steady state line. In the larger model, a small perturbation that either increases or decreases one or many reactant concentrations to disrupt the system, also causes each reactant to move to a different steady state value defined on a steady state curve. In this case, because the system is closed, the small increase or decrease in concentration must be preserved. So the system resets its steady state to account for the change in concentrations. This change in steady state can be tracked with respect to each participating reactant. The resulting changes of steady state act as a stable steady state curve.

## *4.7 Gepasi*

The Gepasi program, outlined in Chapters 2 and 3, was used to verify the results created in Matlab. The benefit of using a simulator such as Gepasi is that it allows for a very friendly user interface, so verifying results is both quick and easy. In this section, the small and large glucose models created and executed in Matlab, will be compared to simulations of these models executed in Gepasi. By evaluating a system using both Matlab and Gepasi, the results can be compared to check for errors in the program.

The Matlab results at 5000 time steps for the small glucose model with kinetic constant values of k4=2, k5=1, k6=2, and k7=1 and initial conditions of 3-Phosphoglycerate = 50, Phosphoglycerate Mutase = 100, Catalyst transition = 0, and 2-Phosphoglycerate = 0.00001 are given by the following.

$$\text{3-Phosphoglycerate} = 0.4563, \tag{4.35}$$

$$\text{Phosphoglycerate Mutase} = 52.2817, \tag{4.36}$$

$$\text{Catalyst transition} = 47.718, \tag{4.37}$$

$$\text{2-Phosphoglycerate} = 1.8254 \tag{4.38}$$

Substituting the same conditions into the Gepasi simulator gives the same values at 5000 time steps.

```
[ Y1] = 4.563564e-001 mM,
[ Y2] = 5.228177e+001 mM,
[ Y3] = 4.771824e+001 mM,
[ Y4] = 1.825425e+000 mM,
```

Although this technique isn't generally needed for small systems, it is valuable for bigger systems. Consider the output after 5 time steps for the large glucose

model #1 listed in Appendix A. The quick key gives a general description of the concentrations the output is describing. For a more detailed understanding, see Appendix A.

```
% ____________ QUICK KEY _____________
%
% CO-ENZYME = y(1), Mg++
% ENZYMES = y(2) - y(11)
% CATALYST TRANSITION = y(12) - y(21)
% ENERGY = y(22) - Y(26)
% SUBSTRATES = y(27) - y(37)
% ____________________________________
ans =
 Columns 1 through 7
 68.4342 999.9999 999.9977 999.9828 999.9132 998.7863 979.5891
 Columns 8 through 14
 80.0914 995.6312 994.6608 993.6994 0.0002 0.0023 0.0172
 Columns 15 through 21
 0.0868 1.2137 20.4109 919.9086 4.3688 5.3393 6.3006
 Columns 22 through 28
 100.4887 973.2848 1.4305 1.4305 978.1586 0.0000 0.0000
 Columns 29 through 35
 0.0000 0.0000 0.0035 0.0005 0.0001 0.0050 0.0001
 Columns 36 through 37
 0.0000 42.2363
```

In this large glucose model, at 5 seconds, the Matlab results are that ATP is y(22) = 100.4887 and ADP is y(23) = 973.2848. This relates quite well to the output from the same model written in Gepasi. The output from Gepasi, given next indicates that at 5 seconds, ATP = 100.19 and ADP = 973.5. The small amount of

error comes in part from the different numerical analysis techniques used in the two systems.

```
RESULTS OF INTEGRATION (after 5.00e+000 s)
[ alphaDg] initial = 5.000000e+002 mM, final = 7.673559e-012 mM
[ ATP] initial = 1.000000e+003 mM, final = 1.001906e+002 mM
[ ENZ1] initial = 1.000000e+003 mM, final = 6.832746e+001 mM
[ ENZ2] initial = 1.000000e+003 mM, final = 9.999999e+002 mM
[ X1] initial = 1.000000e-005 mM, final = 1.575943e-004 mM
[ alphaDg6p] initial = 1.000000e-005 mM, final = 9.488750e-007 mM
[ ADP] initial = 1.000000e+003 mM, final = 9.735061e+002 mM
[ ENZ3] initial = 1.000000e+003 mM, final = 9.999976e+002 mM
[ X2] initial = 1.000000e-005 mM, final = 2.371148e-003 mM
[ BataDf6p] initial = 1.000000e-005 mM, final = 1.208948e-009 mM
[ ENZ4] initial = 1.000000e+003 mM, final = 9.999823e+002 mM
[ X3] initial = 1.000000e-005 mM, final = 1.771586e-002 mM
[ BataDf16b] initial = 1.000000e-005 mM, final = 4.742581e-005 mM
[ ENZ5] initial = 1.000000e+003 mM, final = 9.999110e+002 mM
[ X4] initial = 1.000000e-005 mM, final = 8.900037e-002 mM
[ G3p] initial = 1.000000e-005 mM, final = 3.455546e-003 mM
[ Dihp] initial = 1.000000e-005 mM, final = 4.933821e-004 mM
[ ENZ6] initial = 1.000000e+003 mM, final = 9.987871e+002 mM
[ X5] initial = 1.000000e-005 mM, final = 1.212868e+000 mM
[ NAD] initial = 1.000000e+003 mM, final = 1.435234e+000 mM
[ P] initial = 1.000000e+003 mM, final = 1.435234e+000 mM
[ ENZ7] initial = 1.000000e+003 mM, final = 9.795344e+002 mM
[ X6] initial = 1.000000e-005 mM, final = 2.046566e+001 mM
[ Bis13] initial = 1.000000e-005 mM, final = 5.796210e-005 mM
[ NADH] initial = 1.000000e-005 mM, final = 9.780991e+002 mM
```

```
[ ENZ8] initial = 1.000000e+003 mM, final = 8.004378e+001 mM
[ X7] initial = 1.000000e-005 mM, final = 9.199562e+002 mM
[ Phos3] initial = 1.000000e-005 mM, final = 5.017286e-003 mM
[ ENZ9] initial = 1.000000e+003 mM, final = 9.956060e+002 mM
[ X8] initial = 1.000000e-005 mM, final = 4.394014e+000 mM
[ Phos2] initial = 1.000000e-005 mM, final = 9.054872e-005 mM
[ ENZ10] initial = 1.000000e+003 mM, final = 9.946307e+002 mM
[ X9] initial = 1.000000e-005 mM, final = 5.369277e+000 mM
[ Phos] initial = 1.000000e-005 mM, final = 1.131147e-007 mM
[ ENZ11] initial = 1.000000e+003 mM, final = 9.936708e+002 mM
[ X10] initial = 1.000000e-005 mM, final = 6.329212e+000 mM
[ Pyruvate] initial = 1.000000e-005 mM, final = 4.204531e+001 mM
```

By performing an independent calculation, a numerical algorithm can be partially validated. The Gepasi simulator was a great asset in getting a better understanding of how the larger system should behave.

### *4.8 Summary*

Chapter 4 gave insight into how small perturbations in model parameters and variables affect the system as a whole. Perturbations on the linear model showed variations of steady state behavior within a linear system. Perturbations on the Brusselator showed problems associated with analyzing a linearized model. Perturbations of both reactant concentrations and constant parameters in the Schnackenberg model showed great sensitivity in the form of large amplification that may have broad affects on other systems within a larger model. Hasty's bistable system showed how different magnitudes of perturbations of reactant concentrations can permanently change the behavior of a system. The small Glucose model extended stable steady states points to stable steady state curves and showed how perturbations in reactant concentrations can lead to different steady states within the curve,

while perturbations within parameters can change the shape of the curve. The larger Glucose model then extended the analysis of small systems. Incorporating the small glucose model into the larger Glucose model leads to insight on how to incorporate small systems into larger ones. Furthermore, it shows that doing so can lead to a better understanding of the larger system. Finally, the glucose models were applied to the model simulator Gepasi as an example of how this analysis can be verified using user friendly simulators available over the internet.

# V. Summary and Conclusion

## 5.1 Summary

The goal of this thesis was to understand the underlying behavior of small models and their relation to larger systems in an effort to better understand how the Air Force should construct intracellular models to assist in future toxicology studies. The approach taken in this thesis was to create a system of reaction equations, and then transform that system into a system of differential equations that could be solved. Several models were derived using this technique and then analyzed. New models were created including two models of transcription and translation, a large model of glycolysis, and a small component model of the large glucose model. Then, a number of the previously derived models as well as additional models were evaluated using a sensitivity analysis in an effort to identify how their behavior may change with small perturbations in both initial conditions and constant parameters. This analysis technique was applied to the newly created small glycolysis model and incorporated into the larger model in an effort to understand how the behavior of smaller models affects larger systems.

## 5.2 Conclusion

Reaching the goal of this thesis has led to a number of conclusions. First, the technique of modeling a physical system using biochemical reaction equations and then converting the reaction equations to a system of ordinary differential equations is appropriate for smaller systems. Second, analysis of models using the technique of linearizing about steady state and then determining the stability of the linear model shows that it is capable of displaying interesting behavior in small systems. Third, this technique can be used to analyze larger systems.

The technique of using reaction equations to model transcription, translation, and metabolism was a result of reviewing the modeling techniques presented by

Hasty et al. [14], [13], Schnackenberg [30], and Prigogine et al.[27]. By deriving results concerning existing models that involved reaction equations, it was hoped that more understanding would be gained as to how to develop new models.

Analyzing a model involved starting with the reaction equations and transforming them into a system of ordinary differential equations that were then used to determine the steady state behavior of the system. Examples of steady state behavior in the form of both stable behavior and unstable behavior were evaluated by deriving linear systems from nonlinear systems using linearization. Hypothetical systems including the Brusselator equation and Schnackenbergs equation were linearized and then the behavior near the steady state was determined from the eigenvalues of the linearized system. A small system with multiple steady state points was evaluated to determine how to find the steady states of the system. This knowledge was incorporated in solving for the steady state points of two real-life cellular systems developed by Hasty et al. Using the information gathered from evaluating these systems, two small reaction models involving transcription and translation were presented and derived. A third reaction model involving the glycolysis pathway was created in the form of a system of 35 ordinary differential equations and then solved using a Matlab script.

To discover how concentrations of reactants behave under uncertain conditions, some of the derived models were analyzed graphically to trace the behavior of the systems. The linear system was evaluated first. Results showed behavior as it relates to stable and unstable steady state points. This analysis was expanded to nonlinear systems by evaluating the Brusselator system. Under different conditions, the Busselator showed both convergence to a limit cycle and oscillatory convergence to a steady state point. This analysis was then used to test the variability of various models with respect to changing initial conditions and uncertain constant parameters. By analyzing the Schnackenberg model, under changing initial conditions with one set of parameters, it was shown that the convergence to a limit cycle might be

more dramatic when the initial conditions fall outside the limit cycle. Changing the constant parameters slightly showed great variability in how the model behaved as well. A small change in constant parameters gave rise to oscillatory stability that showed great sensitivity to initial conditions. A slightly larger change in constant parameters created straight line stability. Next, by considering Hasty et al.'s model, small changes in initial conditions led concentrations to move to different stable steady state points. A steady state analysis was performed on Hasty et al.'s model in the form of a bifurcation diagram to determine steady state behavior with respect to uncertain constant parameters. With the knowledge gained from evaluating these models, a small subcomponent of the newly created glucose model was evaluated. After reducing the system from four equations to three, evaluation showed it had multi-stability in the form of a stable steady state curve for every variation of parameters and initial conditions within reasonable model boundaries. Further reduction to a system of two equations removed a free parameter and allowed the reduced system to be evaluated for a single steady state point.

In an effort to understand how small models relate to larger models, the small glucose model was incorporated into the larger glycolysis model. The small system was evaluated under various conditions of the larger model. The evaluation examined changes that took place in the smaller model and how it was affected by conditions imposed by the larger system.

Altogether, the information gathered within this thesis presents a good basis for understanding how to evaluate cellular systems. It also provides an introduction for how the Air Force should construct intracellular models to assist in future toxicology studies.

### *5.3 Recommendations*

Further research into the construction and evaluation of cellular models would benefit Air Force toxicology studies. It remains to be determined how many types

of behavior, such as limit cycle convergence, are attained. Continued analysis of the models described in this thesis, as well as other models, should lead to understanding of how these types of behavior come about.

Continued analysis can be performed by evaluating Hasty et al.'s [14] model without the assumption of fast reactions. Hasty et al.'s model might also be extended to include promoter molecules and evaluated to see how the behavior changes. Also, new models #1 and #2 can be evaluated numerically to determine whether or not they display interesting behavior. Further insight into how to build models displaying interesting behavior may be gained by incorporating models displaying different types of behavior into a larger model. For example, by incorporating a sensitive model, such as the Schnackenberg model, into a system that involves multi-stability, such as Hasty et al.'s model, a small perturbation of reactant concentration in the sensitive system may be amplified in a way that makes reactant concentrations of the multi-stable model move to another stable steady state. Also, by coupling the model described by Chen et al. [7] with the glucose model, transcription, translation, and metabolism can be tied together. Chen's model can be used to regulate the 10 enzymes of the gluoce model thereby controling the prodution of ATP. Also, techniques for solving for the steady state of larger systems should be explored. This might be done by comparing the speed of numerical techniques with simply running the program to steady state.

Models not included in this thesis that should be considered for future evaluation include stochastic models and time delay models. Finally, hypothetical cellular models involving toxicology issues should be created, separated into smaller models, analyzed, and numerically evaluated to give insight into what types of behavior should be studied in greater detail.

# *Appendix A. The Large Glucose Model*

Two models of Glucose are included here. The first is a user friendly version that allows the user to understand and track all the reactants within the model. Furthermore, the output is displayed in an easy to read manner and all inputs into the program can be made easily. This is an appropriate model to use when viewing the system as a whole. In this case, it is used to evaluate the changes in ATP, ADP, and the intermediate concentrations. Due to Matlabs difficult programming nature, evaluating individual aspects of the system requires a more complicated version of the program. This more complex version is included as a way to view individual components of the model, and make changing the system an easier task. In this version, the last equation removed as a way to test the small Glucose model in a larger system under conditions of steady state.

For the convenience of the user, each large Glucose model was written as two Matlab programs. The first program displays the output and allows the user to change initial conditions. It then calls the next program. The second program contains the large Glucose model. All the detail about these programs is included as comments in the programs.

## *A.1 Large Glucose Model #1*

### *A.1.1 Program 1.*

```
% y0 is the initial concentrations of the system.
y0=[1000; 1000; 1000; 1000; 1000; 1000; 1000; 1000; 1000; 1000; 1000;
 .00001; .00001; .00001; .00001; .00001; .00001; .00001; .00001;
 .00001; .00001; 1000; 1000; 1000; 1000; .00001; 500; .00001;
 .00001; .00001; .00001; .00001; .00001; .00001; .00001; .00001;
 .00001];
```

```
tsp=[0,5] % this is the time interval for the calculation
[t,y]=ode23s('Glucose',tsp,y0); % Here the Glucose program is called

n=max(size(y));
y(n,:) % prints out the concentration at the final time step

% Evaluating ATP and ADP concentrations
yT = y(:,12) + y(:,14) + y(:,18) + y(:,21) + y(:,22) + y(:,23);
yp = [y(:,12) y(:,14) y(:,18) y(:,21) y(:,22) y(:,23) yT]
plot(t,yp)

% In the plot, ATP = y(22) and ADP = y(23) concentration changes
% involve intermediate transition steps y(12), y(14), y(18), and
% y(21). The total concentration of these values though time is
% equal to the initial concentrations. Notice that y(12) uses up
% almost all ATP immediately, then gets almost all used up by the
% time y(21) is being created.
```

### *A.1.2 Program 2.*

```
function [dydt] = dydt(t,y)
n=max(size(y));
dydt=ones(37,1);
% There are 37 different molecules that are changing in this system.
% Each reaction has a reaction rate coefficient associated with it.
% The reaction rate coefficients are given in this code by k1-k37.
% The forward reaction rate coefficients are 3 and the backwards
% reaction rate coefficients are 1. In this model, all reaction
% rate coefficients are fixed. In future models, they may be
```

```
% considered variables or functions. NOTE: If there is no backward
% reaction, then the reaction is a one-way reaction.
k1 = 3; %REACTION #1
k2 = 1;
k3 = 3; %REACTION #2
k4 = 3; %REACTION #3
k5 = 1;
k6 = 3; %REACTION #4
k7 = 1;
k8 = 3; %REACTION #5
k9 = 1;
k10 = 3; %REACTION #6
k11 = 3; %REACTION #7
k12 = 1;
k13 = 3; %REACTION #8
k14 = 1;
k15 = 3; %REACTION #9
k16 = 1;
k17 = 3; %REACTION #10
k18 = 1;
k19 = 3; %REACTION #11
k20 = 1;
k21 = 3; %REACTION #12
k22 = 1;
k23 = 3; %REACTION #13
k24 = 1;
k25 = 3; %REACTION #14
k26 = 1;
```

```
k27 = 3; %REACTION #15
k28 = 1;
k29 = 3; %REACTION #16
k30 = 1;
k31 = 3; %REACTION #17
k32 = 1;
k33 = 3; %REACTION #18
k34 = 1;
k35 = 3; %REACTION #19
k36 = 1;
k37 = 3; %REACTION #20

% Each reaction is dependent the molecules formed by previous
% reactions. Therefore, all reaction in this system are
% interconnected through time. This interconnection makes modeling
% the system as a whole, a very difficult process. However, first
% looking at each reaction individually, then incorporating
% the equations for the individual reactions into the complete model
% makes this model easier to construct.

% There are 10 individual reactions in the very simple biochemical
% model of the metabolic pathway, Glycolysis. However, each
% reaction is a catalytic reaction. This being the case, each
% reaction can be split into two reactions that more accurately
% describe the catalytic reaction. SEE: Mathematics Applied to
% Deterministic Problems in the Natural Sciences by Lin, Segel,
% and Handleman. This changes the 10 reactions into 20
% slightly more complex reactions involving hypothetical molecular
```

```
% states of catalytic transition.  Each of the 20 reactions has a
% reaction rate depending on the concentration of molecules entering
% into the system and the concentration of those leaving. Before
% looking at the system as a whole, the individual reaction systems
% are constructed here.
% _____________ QUICK KEY _____________
% CO-ENZYME = y(1), Mg++
% ENZYMES = y(2) - y(11)
% CATALYST TRANSITION = y(12) - y(21)
% ENERGY = y(22) - Y(26)
% SUBSTRATES = y(27) - y(37)
% ____________________________________

% Mg++ + Hexokinase +
% alpha-D-Glucose + ATP (k2)<= =>(k1) Catalyst_transition_1
% is equivalent to
eq1 = -k1*y(27)*y(22)*y(1)*y(2) + k2*y(12);

%                                Mg++ + Hexokinase +
% Catalyst_transition_1 =>(k3) alpha-D-Glucose-6-phosphate + ADP
% is equivalent to
eq2 = -k3*y(12);

% Phosphoglucoisomerase +
% alpha-D-Glucose-6-phosphate (k5)<= =>(k4) Catalyst_transition_2
% is equivalent to
eq3 = -k4*y(28)*y(3) + k5*y(13);
```

```
%                                        Phosphoglucoisomerase +
% Catalyst_transition_2 (k7)<= =>(k6) beta-D-Fructose-6-phosphate
% is equivalent to
eq4 = -k6*y(13) + k7*y(29)*y(3);

% Mg++ + Phosphofructokinase +
% beta-D-Fructose-6-phosphate + ATP (k9)<= =>(k8)
                                                    Catalyst_transition_3
% is equivalent to
eq5 = -k8*y(29)*y(22)*y(1)*y(4) + k9*y(14);

%                                  Mg++ + Phosphofructokinase +
% Catalyst_transition_3 =>(k10) beta-D-Fructose-1,6-bisphosphate + ADP
% is equivalent to
eq6 = -k10*y(14);

% Aldolase +
% beta-D-Fructose-1,6-bisphosphate (k12)<= =>(k11)
                                                     Catalyst_transition_4
% is equivalent to
eq7 = -k11*y(30)*y(5) + k12*y(15);

%                                          Aldolase +
% Catalyst_transition_4 (k14)<= =>(k13) Glyceraldehyde-3-phoshate
                                            + Dihydroxyacetone phosphate
% is equivalent to
eq8 = -k13*y(15) + k14*y(31)*y(32)*y(5);
```

```
% Triose phosphate isomerase +
% Dihydroxyacetone phosphate (k16)<= =>(k15) Catalyst_transition_5
% is equivalent to
eq9 = -k15*y(32)*y(6) + k16*y(16);

%                                        Triose phosphate isomerase +
% Catalyst_transition_5 (k18)<= =>(k17) Glyceraldehyde-3-phoshate
% is equivalent to
eq10 = -k17*y(16) + k18*y(31)*y(6);

% Glyceraldehyde-3-phosphate-dehydrogenase +
% (NAD+ + Pi) + Glyceraldehyde-3-phoshate (k20)<= =>(k19)
                                                  Catalyst_transition_6
% is equivalent to
eq11 = -k19*y(31)*y(25)*y(24)*y(7) + k20*y(17);

%                             Glyceraldehyde-3-phosphate-dehydrogenase +
% Catalyst_transition_6 (k22)<= =>(k21) NADH + 1,3-Bisphosphoglycerate
% is equivalent to
eq12 = -k21*y(17) + k22*y(33)*y(26)*y(7);

% Mg++ + Phosphoglycerate_kinase +
% 1,3-Bisphosphoglycerate + ADP (k24)<= =>(k23) Catalyst_transition_7
% is equivalent to
eq13 = -k23*y(33)*y(23)*y(1)*y(8) + k24*y(18);

%                                       Mg++ + Phosphoglycerate_kinase +
% Catalyst_transition_7 (k26)<= =>(k25) 3-Phosphoglycerate + ATP
```

```
% is equivalent to
eq14 = -k25*y(18) + k26*y(34)*y(22)*y(1)*y(8);

% Phosphoglycerate_mutase +
% 3-Phosphoglycerate (k28) <= => (k27) Catalyst_transition_8
% is equivalent to
eq15 = -k27*y(34)*y(9) + k28*y(19);

%                                        Phosphoglycerate_mutase +
% Catalyst_transition_7 (k30)<= =>(k29) 2-Phosphoglycerate
% is equivalent to
eq16 = -k29*y(19) + k30*y(35)*y(9);

% Mg++ + Enolase +
% 2-Phosphoglycerate (k32)<= =>(k31) Catalyst_transition_9
% is equivalent to
eq17 = -k31*y(35)*y(1)*y(10) + k32*y(20);

%                                        Mg++ + Enolase +
% Catalyst_transition_9 (k34)<= =>(k33) Phophoenolpyruvate
% is equivalent to
eq18 = -k33*y(20) + k34*y(36)*y(1)*y(10);

% Mg++ + Pyruvate_Kinase +
% Phophoenolpyruvate + ADP (k36)<= =>(k35) Catalyst_transition_10
% is equivalent to
eq19 = -k35*y(36)*y(23)*y(1)*y(11) + k36*y(21);
```

```
%                                    Mg++ + Pyruvate_Kinase +
% Catalyst_transition_10 =>(k37) Pyruvate + ATP
% is equivalent to
eq20 = -k37*y(21);

% Now each of the 37 molecules change with respect to time.
% Each of the equations is the reaction rate
% written with respect to a single molecule on the left hand
% side of the biochemical equation. The negative of that
% equation will give the reaction rate written with respect
% to a single molecule on the right hand side of the biochemical
% equation. To find the reaction rate of a specific molecule
% over the entire system, add each equation where the molecule
% appears on the left hand side of the biochemical equation.
% Of course, if the molecule appears on the right hand side
% of the biochemical equation, add the negative of the above
% equations.

dydt(1) = eq1 - eq2 + eq5 - eq6 + eq13 - eq14
          + eq17 - eq18 + eq19 - eq20;
% Change of Mg++ with respect to time
dydt(2) = eq1 - eq2;
% Change of Hexokinase with respect to time
dydt(3) = eq3 - eq4;
% Change of Phosphoglucoisomerase with respect to time
dydt(4) = eq5 - eq6;
% Change of Phosphofructokinase with respect to time
dydt(5) = eq7 - eq8;
```

```
% Change of Aldolase with respect to time
dydt(6) = eq9 - eq10;
% Change of Triose phosphate isomerase with respect to time
dydt(7) = eq11 - eq12;
% Change of Glyceraldehyde-3-phosphate-dehydrogenase
% with respect to time
dydt(8) = eq13 - eq14;
% Change of Phosphoglycerate kinase with respect to time
dydt(9) = eq15 - eq16;
% Change of Phosphoglycerate mutase with respect to time
dydt(10) = eq17 - eq18;
% Change of Enolase with respect to time
dydt(11) = eq19 - eq20;
% Change of Pyruvate kinase with respect to time
dydt(12) = -eq1 + eq2;
% Change of Catalyst_transition_1 with respect to time
dydt(13) = -eq3 + eq4;
% Change of Catalyst_transition_2 with respect to time
dydt(14) = -eq5 + eq6;
% Change of Catalyst_transition_3 with respect to time
dydt(15) = -eq7 + eq8;
% Change of Catalyst_transition_4 with respect to time
dydt(16) = -eq9 + eq10;
% Change of Catalyst_transition_5 with respect to time
dydt(17) = -eq11 + eq12;
% Change of Catalyst_transition_6 with respect to time
dydt(18) = -eq13 + eq14;
% Change of Catalyst_transition_7 with respect to time
```

```
dydt(19) = -eq15 + eq16;
% Change of Catalyst_transition_8 with respect to time
dydt(20) = -eq17 + eq18;
% Change of Catalyst_transition_9 with respect to time
dydt(21) = -eq19 + eq20;
% Change of Catalyst_transition_10 with respect to time
dydt(22) = eq1 + eq5 - eq14 - eq20;
% Change of ATP with respect to time
dydt(23) = -eq2 - eq6 + eq13 + eq19;
% Change of ADP with respect to time
dydt(24) = eq11;
% Change of NAD+ with respect to time
dydt(25) = eq11;
% Change of Pi with respect to time
dydt(26) = -eq12;
% Change of NADH with respect to time
dydt(27) = eq1;
% Change of alpha-D-Glucose with respect to time
dydt(28) = -eq2 + eq3;
% Change of alpha-D-Glucose-6-phosphate with respect to time
dydt(29) = -eq4 + eq5;
% Change of beta-D-Fructose-6-phosphate with respect to time
dydt(30) = -eq6 + eq7;
% Change of beta-D-Fructose-1,6-bisphosphate with respect to time
dydt(31) = -eq8 - eq10 + eq11;
% Change of Glyceraldehyde-3-phoshate with respect to time
dydt(32) = -eq8 + eq9;
% Change of Dihydroxyacetone phosphate with respect to time
```

```
dydt(33) = -eq12 + eq13;
% Change of 1,3-Bisphosphoglycerate with respect to time
dydt(34) = -eq14 + eq15;
% Change of 3-Phosphoglycerate with respect to time
dydt(35) = -eq16 + eq17;
% Change of 2-Phosphoglycerate with respect to time
dydt(36) = -eq18 + eq19;
% Change of Phophoenolpyruvate with respect to time
dydt(37) = -eq20;
% Change of Pyruvate with respect to time
```

Figure A.1 is a graph that describes how ATP and ADP move through the larger model. This larger model is simply an example of how small systems can be incorporated into larger systems. In this example, at time $t = 0$, ATP has a concentration of 1000 units, ADP has a concentration of 1000 units, and the four intermediate steps have a concentration of .0001 units.

It is clear that throughout this system, for the first five time steps, ATP, ADP, and the intermediate steps change concentrations quite rapidly. However, it is also clear by the line that represents the total concentrations of all the reactants, that no reactants diverge. This is a property that held true for the smaller system and is also true for the larger system.

As time moves out to 5000 time steps, it is clear that the reactants still do not diverge. Stable systems such as these are what biological systems rely on to survive. Through these systems, they may control their to and begin to settle down into a form that

## *A.2 Large Glucose Model #2*

### *A.2.1 Program 1.*

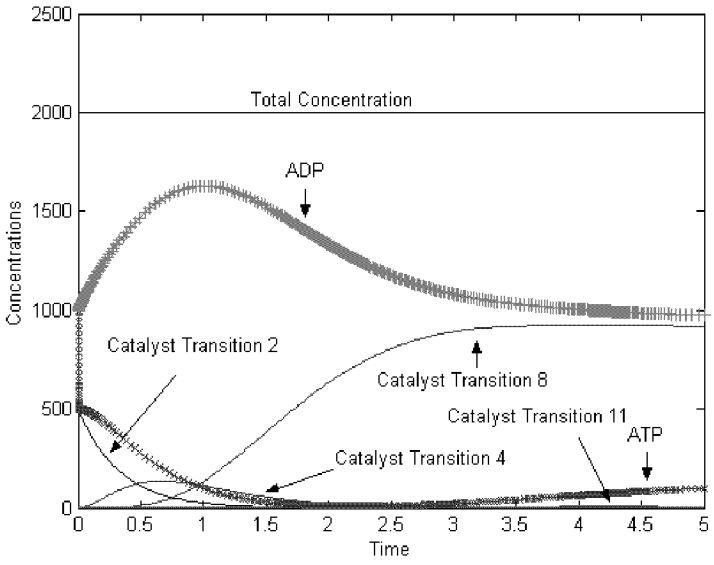


Figure A.1. Concentration graph of ATP, ADP and associated catalyst transitions in the large glucose model

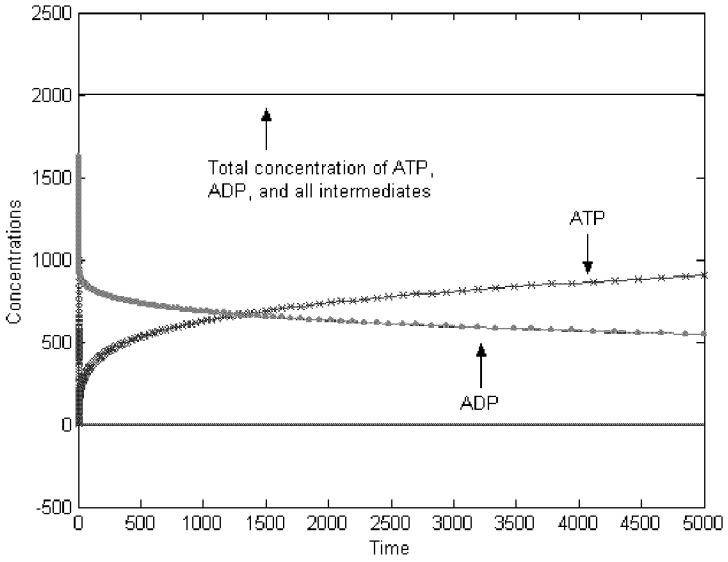


Figure A.2. Concentration graph of ATP, ADP and associated catalyst transitions in the large glucose model

```
% Initial concentrations
y0=[500; 10; 1000; 100; 1000; .00001; .00001; 1000; .00001; .00001;
 1000; .00001; .00001; 1000; .00001; .00001; .00001; 1000; .00001;
 1000; 1000; .00001; 1000; .00001; .00001; 10; .00001; 100.00001;
 1000; .00001; .00001; 10; .00001; .00001];

tsp=[0,5000] % plug in the time interval
[t,y]=ode23s('Glucose_tapered',tsp,y0);
m=max(size(y));

% Evaluating reaction concentrations
c= y(:,29) + y(:,30);
yp = [y(:,28) y(:,29) y(:,30) y(:,31), c]

% plots the concentrations of each reactant in reaction 8
% with respect to time

%plot(t,yp)
k31 = 1.1; %REACTION #17
k32 = 1;
k33 = 1.1; %REACTION #18
k34 = 1;
c0 = c(1,1);

step = m; % Input the time step (an integer between 1 and n)
alpha = y(step,29);
time_step = t(step);
%
```

```
% Checks the steady state value
y_28=(c0*k32 - alpha*k32)/(alpha*k31);
y_31=(c0*k33 - alpha*k33)/(alpha*k34);
[y(m,:), t(m)] % prints out the concentration at the final time step
steady_state = [time_step y_28 alpha c0-alpha y_31]
%
% plots how the concentrations of the reactants move through time.
% plot3(y(:,28), y(:,29), y(:,31))
```

### *A.2.2 Program 2.*

```
function [dydt] = dydt(t,y)
n=max(size(y));
dydt=ones(34,1);
% This is the number of elements
% in this pathway that go through
% a change over time.
% These are the reaction rate coefficients
% The reaction rates are taking place with
% rate coefficient concentrations of millimoles/sec.

k1 = 3; %REACTION #1
k2 = .01;
k3 = 3; %REACTION #2
k4 = 3; %REACTION #3
k5 = .01;
k6 = 3; %REACTION #4
k7 = .01;
k8 = 3; %REACTION #5
```

```
k9 = .01;
k10 = 3; %REACTION #6
k11 = 3; %REACTION #7
k12 = .01;
k13 = 3; %REACTION #8
k14 = .01;
k15 = 3; %REACTION #9
k16 = .01;
k17 = 3; %REACTION #10
k18 = .01;
k19 = 3; %REACTION #11
k20 = .01;
k21 = 3; %REACTION #12
k22 = .01;
k23 = 3; %REACTION #13
k24 = .01;
k25 = 30; %REACTION #14
k26 = .0001;
k27 = 1.1; %REACTION #15
k28 = 1;
k29 = 3; %REACTION #16
k30 = .01;
k31 = 1.1; %REACTION #17
k32 = 1;
k33 = 1.1; %REACTION #18
k34 = 1;
% k35 = 3; %REACTION #19 %commented out
% k36 = 1;                %commented out
```

```
% k37 = 3; %REACTION #20 %commented out


A = -k1*y(1)*y(2)*y(4)*y(5) + k2*y(6);
B = -k3*y(6);
C = -k4*y(7)*y(8) + k5*y(9);
D = -k6*y(9) + k7*y(8)*y(10);
E = -k8*y(10)*y(2)*y(4)*y(11) + k9*y(12);
F = -k10*y(12);
G = -k11*y(13)*y(14) + k12*y(15);
H = -k13*y(15) + k14*y(14)*y(16)*y(17);
I = -k15*y(17)*y(18) + k16*y(19);
J = -k17*y(19) + k18*y(18)*y(16);
K = -k19*y(16)*y(20)*y(21)*y(23) + k20*y(24);
L = -k21*y(24) + k22*y(22)*y(23)*y(25);
M = -k23*y(25)*y(3)*y(4)*y(26) + k24*y(27);
N = -k25*y(27) + k26*y(2)*y(4)*y(26)*y(28);
O = -k27*y(28)*y(29) + k28*y(30);
P = -k29*y(30) + k30*y(29)*y(31);
Q = -k31*y(31)*y(4)*y(32) + k32*y(33);
R = -k33*y(33) + k34*y(4)*y(32)*y(34);
% S = -k35*y(34)*y(3)*y(4)*y(35) + k36*y(36);   %commented out
% T = -k37*y(36);                               %commented out
dydt(1) = A;
dydt(2) = A + E - N; % - T;    %values commented out
dydt(3) = -B -F + M; % + S;    %values commented out
dydt(4) = A - B + E - F + M - N + Q - R; % + S - T;
                         % values of dydt(4) are commented out
dydt(5) = A - B;
```

```
dydt(6) = -A + B;
dydt(7) = -B + C;
dydt(8) = C - D;
dydt(9) = -C + D;
dydt(10) = -D + E;
dydt(11) = E - F;
dydt(12) = -E + F;
dydt(13) = -F + G;
dydt(14) = G - H;
dydt(15) = -G + H;
dydt(16) = -H - J + K;
dydt(17) = -H + I;
dydt(18) = I - J;
dydt(19) = -I + J;
dydt(20) = K;
dydt(21) = K;
dydt(22) = -L;
dydt(23) = K - L;
dydt(24) = -K + L;
dydt(25) = -L + M;
dydt(26) = M - N;
dydt(27) = -M + N;
dydt(28) = -N + O;
dydt(29) = O - P;
dydt(30) = -O + P;
dydt(31) = -P + Q;
dydt(32) = Q - R;
dydt(33) = -Q + R;
```

```
dydt(34) = -R; % + S; %values commented out
% dydt(35) = S - T;   %commented out
% dydt(36) = -S + T;  %commented out
% dydt(37) = -T;      %commented out
```

## *Bibliography*


1. Alberts, B.; Bray, D.; Lewis J.; Raff M.; Roberts K.; Watson J.D. *Molecular Biology of The Cell, Third Edition.* New York: Garland Publishing, Inc.

2. Atkinson, K.E. *An introduction to Numerical Analysis.* New York: John Wiley and Sons, 1978.

3. Baras, F.; Mansour, M.M.; and J.E. Pearson. "Microscopic Simulation of Chemical Bistability in Homogeneous Systems," *Journal of Chemical Physics*, *105*(18):8257–8261 (November 1996).

4. Bhalla, U.S.; and R. Iyengar. "Emergent Properties of Networks of Biological Signaling Pathways," *Science*, *283*:381–387 (January 1999).

5. Brauer, F. and A.N. Nohel. *The Qualitative Theory of Ordinary Differential Equations an Introduction.* New York: Dover Publications Inc., 1969.

6. Campbell, N.A., Reece J.B. Mitchell L.G. *Biology.* New York: Addison Wesley Longman, Inc., 1999.

7. Chen, T.; He, H.L.; and G.M. Church. "Modeling Gene Expression With Differential Equations," *Pacific Symposium on Biocomputing*, *4*:29–40 (1999). http://psb.stanford.edu/psb-online.

8. Cobelli, C.; Audoly, S.; Bellu G.; D'Angio L.; and M.P. Saccomani. "Global Identifiability of Nonlinear Models of Biological Systems," *IEEE Transactions on Biomedical Engineering*, *vol. 48*(1):55–65 (January 2001).

9. Gillespie, D.T. "Exact Stochastic Simulation of Coupled Chemical Reactions,"

10. Glick, B.R.; and J.J. Pasternak. *Molecular Biotechnology.* 1325 Massachusetts Aveenue NW, Washington, DC 20005-4171: American Society for Microbiology, 1998.

11. Gray, P. and S. Scott. *Chemical Oscillations and Instabilities.* New York: Oxford Science Publications, Oxford University Press, 1990.

12. Hale, J. and H. Kocak. *Dynamics and Bifurcations.* New York: Springer-Verlag, 1991.

13. Hasty, J.; Pradines, J.; Dolnik M.; and J.J. Collins. "Noise-Based Switches and Amplifiers for Gene Expression," *The National Academy of Sciences of the USA*, *97*(5):2075–2080 (February 2000).

14. Hasty, J.; Pradines, J.; Dolnik M.; and J.J. Collins. "Stochastic Regulation of Gene Expression," *American Institute of Physics*, 191–196 (2000).

15. Heinrich, R.; Rapoport, S.M.; and T.A. Rapoport. "Metabolic Regulation and Mathematical Models," *Progress in Biophysics and Molecular Biology*, *32*:1–82 (1977).

16. Johnson, A.D., Poteete A.R. Lauer G. Sauer R.T. Ackers G.K. and M. Ptashne. "Lambda Repressor and Cro - Components of an efficient molecular switch," *Nature, London*, *294*:217–223 (November 1981).

17. Lewin, B. *Genes*. New York: Oxford University Press, 1997.

18. Lin, C.C.; Segel, L.A.; and G.H. Handelman. *Mathematics Applied to Deterministic Problems in the Natural Sciences*. Philadelphia: SIAM.

19. Maloy, S.R.; Cronan Jr., J.E. Jr.; Freifelder D. *Microbial Genetics*. Boston: Jones and Bartlett Publishers, 1994.

20. McAdams, H.H.; and A. Arkin. "Simulation of Prokaryotic Genetic Circuits," *Annual Review Biophys. Biomol. Struct.*, *27*:199–224 (1998).

21. McAdams, H.H.; and A. Arkin. "It's a Noisy Business! Genetic Regulation at the Nanomolar Scale," *T.I.G.*, *15*(2):65–69 (February 1999).

22. McAdams, H.H.; and L. Shapiro. "Circuit Simulation of Genetic Networks," *Science*, *269*:650–656 (August 1995).

23. Mendes, P. "GEPASI: a software package for modelling the dynamics, steady states and control of biochemical and other systems," *Cabios*, *9*(5):563–571 (1993).

24. Meyer, T., Waeber G. Lin J. Beckmann W. Habener J. "The promoter of the gene encoding cAMP response element binding protein contains cAMP response elements: evidence for positive autoregulation of gene transcription," *Endocrinology*, *132*:770–780 (1993).

25. Personal Communication, Dr. Dennis Quinn and Dr. Brent Foy, "Funded DAGSI Proposal," December 2000.

26. Petzold, L. "Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations.," *SIAM Journal of Science, Statistics, and Computers*, *4*:136–148 (1983).

27. Prigogine, I. and R. Lefever. "Symmetry Breaking Instabilities in Dissipative Systems II," *The Journal of Chemical Physics*, *48*(4):1695–1700 (February 1968).

28. Ptashne, M., Jeffrey A. Johnson A.D. Maurer R. Meyer B.J. Pabo C.O. Roberts T.M. and R.T. Sauer. "How the lambda Repressor and Cro Work," *Cell*, *19*:1–11 (1980).

29. Savageau, M.A. "Rules For the Evoluation of Gene Circuitry," *Pacific Symposium on Biocomputing*, *3*:54–65 (1998). http://psb.stanford.edu/psb-online.

30. Schnakenberg, J. "Simple Chemical Reaction Systems with Limit Cycle Behavior," *Journal of theoretical biology*, *81*:389–400 (1979).

31. Selkov, E.E. "Self-Oscillations in Glycolysis," *European Journal of Biochemestry*, *4*:79–86 (1968).

32. Smolen, P.;Baxter, D.A.;Byrne J.H. "Frequency selectivity, multistability, and oscillations emerge from models of genetic regulatory systems," *The American Physiological Society*, *43*:C531–C542 (1998).

33. Tomita, M. et al. "The E-Cell Project: Towards Integrative Simulation of Cellular Processes," *New Generation Computing* (September 1999).

34. Voet, D.; and J. Voet. *Biochemistry, Second Edition*. One Wiley Drive, Somerset, NJ 08875: John Wiley Sons, Inc., 1995.

# *Vita*

2LT Campbell was born in Minneapolis, Minnesota on June 1977. He grew up in West St. Paul, Minnesota where he graduated from high school in 1996. He attended Montana State University in Bozeman, Montana and graduated with a Bachelor of Science in Mathematics in 2000.

2LT Campbell's first assignment was to attend the Air Force Institute of Technology (AFIT) in residence at Wright-Patterson Air Force Base located in Dayton, Ohio to complete a Masters of Science in Applied Mathematics Analysis.

2LT Campbell married his lovely wife while completing his thesis.

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 074-0188*



| 1. REPORT DATE *(DD-MM-YYYY)* | 2. REPORT TYPE | 3. DATES COVERED *(From – To)* |
|---|---|---|
| 14-03-2002 | **Master's Thesis** | Sept 2001 – Mar 2002 |

**4. TITLE AND SUBTITLE**

CELL MODELING

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**

Campbell, Matthew, L., Second Lieutenant, USAF

**5d. PROJECT NUMBER**
ENR # 2001-049

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)**
Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 P Street, Building 640
WPAFB OH 45433-7765

**8. PERFORMING ORGANIZATION REPORT NUMBER**

AFIT/GAM/ENC/02M-01

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**

AFOSR
Attn: Dr. Kozumbo (703) 696-7720
801 N. Randolph St. Room 732
Arlington, VA 22203

DAGSI
3155 Research Blvd, Suite 205
Kettering, OH 45420
(937) 781-4080

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION/AVAILABILITY STATEMENT**

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**

The Air Force is currently developing new products that incorporate a variety of chemicals which may come in contact with product users. To define which chemicals are dangerous to the user, toxicity studies have been performed. However, analysis of toxicity ultimately requires models of the exposed cellular systems. This thesis provides an introduction of how to model and analyze small and large cellular systems. Understanding the underlying behavior of small models and their relation to large systems will lead to a better understanding of how the Air Force should construct intracellular models to assist in future toxicology studies.

Developing analysis techniques to include steady state analysis through linearization, and then considering small reaction systems, such as the Brusselator and Schnackenberg models, led to a basic understanding of model behavior. This knowledge was applied to create new models in an effort to begin a transition from previously created models to the construction of models unique to the Air Force.

Sensitivity analyses performed on existing systems furthered research efforts by developing knowledge of how systems behave under various initial conditions and perturbations of uncertain constant parameters. Analysis displayed great sensitivity within some models. This analysis was applied to a new model to look for interesting behavior such as oscillatory convergence. The new model was then incorporated into a larger model to determine how its behavior changed with respect to changes in the larger model.

This knowledge of how small systems behave in relation to larger systems should help the Air Force to develop and analyze intracellular toxicology models.

**15. SUBJECT TERMS**

Cell Modeling, Cellular Analysis, Reaction Equations

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | Dr. Dennis W. Quinn (ENC) |
| U | U | U | UU | 139 | **19b. TELEPHONE NUMBER** *(Include area code)* (937) 255-6565, ext 4522; e-mail: Dennis.Quinn@afit.edu |